DOCUMENT RESUME

ED 026 793 EC 003 658

By-Edmonson, Barbara; And Others
Social Inference Training of Retarded Adolescents at the Pre-Vocational Level.
Kansas Univ., Kansas City. Medical Center.
Spons Agency-Vocational Rehabilitation Administration (DHEW), Washington, D.C.
Pub Date 68
Grant-RD-1388-P
Note-171p.
EDRS Price MF-$0.75 HC-$8.65
Descriptors-*Attitudes, *Behavior, Behavior Change, Behavior Rating Scales, Comprehension Development, Educable Mentally Handicapped, *Exceptional Child Research, Institutionalized (Persons), Interpersonal Competence, Junior High School Students, *Mentally Handicapped, Prevocational Education, Social Attitudes, Socially Maladjusted, Student Evaluation, Test Reliability, *Tests, Test Validity
Identifiers-Test of Social Inference, TSI

The relationship between the difficulty with which the retarded adolescent decodes visual social cues and the appropriateness of his social behavior was the basis of a study to determine whether social cue decoding could be taught the retarded in a manner analogous to the teaching of reading and writing, with beneficial behavioral results. The Test of Social Inference (TSI) was developed and used as the before and after criterion of social comprehension. During four trials, 11 classes of retarded adolescents at the prevocational level were provided experimental lessons designed to promote social awareness daily for periods of from 6 to 10 weeks as a part of their special education. Two control classes were provided placebo materials and five were provided contrast materials to supplement their usual curricula, and one control class had no change made in its usual program. All pupils were given social behavioral ratings by their teachers before and after the periods of treatment. The association between social comprehension and social behavior was confirmed by significant correlations between the pupils' TSI scores and their social behavioral ratings (p=.001). Improved comprehension as a result of special training was confirmed by significantly higher TSI scores following treatment (p.05). Improved behavior associated with gains in social comprehension was supported by results only of the initial trial. (Author/RP)

ED026793

SOCIAL INFERENCE TRAINING
OF RETARDED ADOLESCENTS AT THE PRE-VOCATIONAL LEVEL

Barbara Edmonson
Henry Leland
Ethel M. Leach, Project Co-directors

Final report of research and demonstration project RD-1388-P supported in part by the Vocational Rehabilitation Administration of the U.S. Office of Health, Education and Welfare

UNIVERSITY OF KANSAS MEDICAL CENTER
KANSAS CITY, KANSAS 66103

1968

EC003658

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

SOCIAL INFERENCE TRAINING OF RETARDED ADOLESCENTS AT THE PRE-VOCATIONAL LEVEL

Barbara Edmonson
Henry Leland
Ethel Leach
Project Co-Directors

with the assistance of
John E. deJung

University of Kansas Medical Center
Kansas City, Kansas, 66103
1968

Final report of research and demonstration project RD-1388-P supported in part by the Vocational Rehabilitation Administration of the U. S. Office of Health, Education and Welfare and in part by the University of Kansas Medical Center

SIGNIFICANT FINDINGS FOR THE REHABILITATION WORKER

The Test of Social Inference

The Test of Social Inference (TSI), developed as an experimental instrument by this project, provides a social comprehension score whose relevance to the adolescent and adult retardate's social behavioral adequacy is in part indicated by Pearson product moment correlations significant at the .001 level between subjects' TSI scores and ratings given them by teachers with respect to social range, social relationships, peer acceptance, attentiveness and calmness, as defined on on a battery of behavior rating scales.

In further indications of the meaningfulness of the test, the TSI scores of EMR subjects tested were more predictive of the placement of the samples--institutional, public school special class, or public school regular class--than their IQ's. A description of the TSI,with reliability and validity data is to be found on pages 2.12-2.28 of Chapter II and again in Chapter VI.

The Retardate's Deficit in Social Comprehension

Although it was anticipated that the retarded adolescent and young adult would score significantly lower on the TSI than would his nonretarded (NMR) counterpart, it was a surprise to find a wide range of TSI scores at each level of age and IQ of the educable retarded pupil. Within the public school special education classes (PS-EMR) tested, a few EMR pupils earned TSI scores equivalent to those of NMR subjects. At the other extreme, the performance by some suggested virtually no decoding of the social cues normally used in the modulation of one's own behavior. The institutional educable retarded subjects (Inst-EMR) tested had similarly wide ranges of TSI scores at each level of age and IQ, but with an overall deficit significantly greater than that of the PS-EMR. At each level of age and IQ, the TSI score of the Inst-EMR subject was below the TSI score of the PS-EMR subject. Data relevant to the social comprehension deficit is presented on pages 2.13-2.22 of Chapter II and in part 2 of Chapter VI.

Remediability

The experimental trials,in six junior high special education classes and five prevocational classes in state institutions,of lessons prepared so as to illustrate social situations and expected behaviors, resulted in significant gains in social comprehension (gains in TSI scores). Of the 127 pupils in the experimental classes, 87 percent had TSI score gains when retested after use of the lessons; their gains averaging 12 TSI points in comparison with average gains by nonexperimentally treated pupils of 5 TSI points over comparable time intervals.

The gains associated with the lessons should not be considered maximal because of the restrictions placed upon their use during the trials. Teachers were furnished the "canned" materials with little or no opportunity to become familiar with objectives or content. In addition, lessons had to be completed within a stipulated period, allowing of no repetition nor possibility of lesson supplement. Data from the experimental trials are presented in Chapter V. A discussion of factors related to gains is contained in Chapter VI.

Gains in Relation to Age and IQ

TSI gains were found moderately positively correlated with age with the few instances of negative gain being restricted to the lower aged subjects. IQ appeared unrelated to TSI gains. Within the ranges represented by the project's samples, neither age nor IQ appeared restrictive of significant gains in association with educative treatment.

Gains in Relation to Sex

Boys made greater TSI gains than girls in most classes, the difference being particularly marked in the two trials of experimental curriculum II in two state institutions. Because of the unevenness of program effect, the project was extended to permit major curricular revisions, and a further field trial. In this final trial, involving the abridged use of Curriculum III in four junior high special education classes, the gains by girls did not differ significantly from those by boys.

Cause of the Social Comprehension Deficit

The discovery of the wide ranges in social comprehension at each level of age and IQ, in conjunction with the gains associated with brief remedial treatment is important, posing as it does the research question as to the cause of the deficit. Allied to this question is the social comprehension deficit of the institutional subject that is even more marked than that of his community-based peer.

PROJECT STAFF AND CONSULTANTS

Barbara Edmonson, M.A. - Project Co-Director, Principal Investigator
Ethel M. Leach, Ed.D. - " "
Henry Leland, Ph.D. - " "

Gail Achtenberg, B.A. - Research Assistant
Jean Barrington, B.A. - " "
Donna Brahl - " "
Demaris Carrell, M.A. - " "
Florence Goldman, M.A. - " "
Beatrice Gundle, M.A. - " "
Jane Havens, B.S. - " "
Martha Reed, B.A. - " "

Alice J. Harford - Secretary
John Reed - Photographer
Elizabeth Ganser - Artist
Ivan Ganser, M.F.A. - Artist
Elizabeth Galton, B.A. - Research Fellow

Consultants

William Arndt, Ph.D.
Jerry Chaffin, Ed.D.
Jean Clark, M.S.W.
John E. deJung, Ed.D.
Morton Goldman, Ph.D.
Paul Gump, Ph.D.
Sheila M. Jackson, B.Ed.
Louise Luckett, M.Ed.
Julia McHale, Ph.D.
Lucille Morris, M.A.
Kazuo Nihira, Ph.D.
James Payne, M.A.
Charles Solley, Ph.D.
Jeanne Truett, M.A.
Robert Tyrrell, Ph.D.
Herbert Wright, Ph.D.

ACKNOWLEDGMENTS

- For rich information concerning the social behavior and problems of retarded youth, the following counsellors, social workers and vocational trainees should be credited for their descriptions and reports which shaped portions of the lessons contained in Social Perceptual Training for Community Living: Pre-Vocational Units for Retarded Youth.

Carol Campbell	Vocational Rehabilitation Counsellor
Daisy Caress	Director of Social Services, Parsons, Kansas State Hospital
Jerry Chaffin	Principal Investigator, A Selected Demonstration for the Vocational Training of Mentally Retarded Youth in Public High Schools (Kansas Work Study Project)
Jean Clark	(formerly) Chief Social Worker, Vocational Rehabilitation Unit, Topeka, Kansas
David Hale	Director Rehabilitation Services, Goodwill Industries, Kansas City, Mo.
Jack Martin	Field Research Director, Peer Knowledge Among Retarded Youth
Joyce North	Family Service Consultant, Kansas Work-Study Project
Susy Paddock	Rehabilitation Supervisor, Goodwill Industries, Kansas City, Mo.
James Payne	Vocational Counselor, Kansas Work-Study Project
T.L.Satterfield	(formerly) Director,Rehabilitation Services Goodwill Industries, Kansas City, Mo.
Richard Schutz	Director, Vocational Rehabilitation Unit, Topeka, Kansas
Lloyd Williams	Rehabilitation Supervisor, Goodwill Industries, Kansas City, Mo.

- For their cooperation in arranging for trials of the experimental social perceptual lessons much appreciation is due the following administrators of school and institutional educational programs:

Trial 1:	Floyd Hudson (former) Coordinator of Special Education, Shawnee Mission, Kansas, Public High School District. Robert Sullivan, Coordinator of Special Education, Turner, Kansas, Public High School District
Trial 2: (First revision)	Henry Leland, Coordinator of Professional Training Juanita Haywood, Director of Education Parsons, Kansas State Hospital and Training Center
Trial 3 (First revision)	James Borren, Superintendent Kathleen Alexander, Director of Education Hissom Memorial Center, Sand Springs, Oklahoma
Trial 4 (Second revision)	Afton Bridges, Coordinator of Special Education Springfield, Missouri, Public School District James Gray, Coordinator of Special Education Shawnee Mission, Kansas, Public High School District

- For a fifth trial of the units and the opportunity to incorporate portions of his data into the report, appreciation is extended to Gary M. Clark, Ph.D., Assistant Professor, University of Illinois. Clark's independent trial of the first revision of the social perceptual units was undertaken at George F. Peabody College as a doctoral dissertation project (Clark, 1967).

- Most essential to the project were the teachers who used and criticized the experimental materials. By means of their suggestions and criticisms each of the following had some influence upon the completed form of the social perceptual units:

Lowell Alexander
Carlos Bradfield
Frank Brandenberg
Ronald Brooks
Phyllis Conwell
Mildred Howard
Ruby Hughes
Sheila M. Jackson
Wayne Koons
Hazel Martin
Marjory Smith
Janette Stafford
Clementine Steger
Betty Stephens
Three who are anonymous in Nashville, Tennessee

- For their assistance in pre-trial arrangements, thanks are extended to Julia McHale, Ph.D., (formerly) Professor of Psychology, Oklahoma State University, Oklahoma City, and to John Patterson, Consultant, Program for Mental Retardation, State Department of Education, Jefferson City, Mo.

- Grateful acknowledgment should be made of the unflagging help and diverse talents of research assistants Donna Brahl, Demaris Carrell, Beatrice Gundle, Florence Goldman, Jane Havens and Martha Reed.

- In connection with trials of the Test of Social Inference, the assistance of Daryle Bass, Parsons State Hospital psychology trainee, is acknowledged with warm appreciation.

- Credit for a major portion of the project's tactical guidance is due John E. de Jung, former project director, who remained as project consultant following his transfer in 1964 to the University of Oregon.

- As Principal Investigator and co-director throughout, responsibility for the project's conduct and the contents of this report is assumed by the undersigned.

Barbara Edmonson, Principal Investigator
Research Associate in Education
University of Kansas Medical Center
Kansas City, Kansas 66103

TABLE OF CONTENTS

Page
Significant Findings for the Rehabilitation Worker iii
Project Staff and Consultants v
Acknowledgments vi
Table of Contents viii
Table of Tables ix
Table of Figures xiii
I Introduction 1.1
Background 1.1
Project Objective 1.2
Related Research 1.3
Project Overview 1.8
II Test of Social Inference (TSI) 2.1
A. Early Development 2.1
B. Response Classification 2.4
C. Test Description - 1967 Revision 2.7
D. Inter-Scorer Reliability 2.8
E. Test-Retest Reliability 2.8
F. Validity: Sampling,Age, and IQ Differences 2.12
TSI Scores in Relation to Sample Difference 2.12
TSI " " " " Age 2.18
TSI " " " " IQ 2.19
TSI " " " " Behavior Rating 2.22
TSI " " " " Other Criteria 2.26
G. Summary 2.28
III The Rating of Social Behavior 3.1
A. Direct Observational Behavior 3.1
B. Description of Teacher and Outside Rater Behavior Rating Scale 3.3
C. Examination of Teacher Behavior Rating Scales (T-BR) 3.5
D. Examination of Outside Rater Rating Scales (O-BR) 3.18
E. Summary 3.22
IV Social Perceptual Training Materials 4.1
A. Social Cue Training Alternatives 4.1
B. The Social Anthropological Outlook 4.2
C. The Retardate's Behavioral Repertoire 4.4
D. Social Perceptual Training Lessons 4.9
E. Evaluation of Lessons 4.15
V Trials of Remedial Materials and Results 5.1
A. Trial 1 5.2
1. Procedures 5.2
2. Results 5.4
B. Trial 2 5.9
1. Procedures 5.9
2. Results 5.12
C. Trial 3 5.18
1. Procedures 5.18
2. Results 5.20

D. Trial 4 5.24
  1. Procedures 5.24
  2. Results 5.26
E. Summary 5.30
VI Discussion of Results and Implications 6.1

(I) Discussion

A. Test of Social Inference 6.1
B. Social Comprehension Deficit of the Retarded as Measured by the TSI 6.3
C. Effectiveness of the Remedial Programs 6.6
D. Relation of TSI Gains to Gains in Behavioral Ratings 6.9

(II) Implications

A. Implications for Research and Rehabilitation of the Social Comprehension Deficit 6.10
B. Implications for Research and Rehabilitation of the Remedial Gains in Social Comprehension 6.13
VII Overview 7.1
A. Rationale, Goals, Procedural Design 7.1
B. The Test of Social Inference 7.2
C. Social Behavioral Rating Scales 7.2
D. Social Perceptual Training 7.3
E. Experimental Trials 7.4
F. Results and Implications 7.8

TABLE OF TABLES

Table		Page
2.1	Resume of Pilot Trials for Development of TSI	2.2
2.2	Samples of Nonretarded and Retarded Subjects from Whom Test-Retest Scores were Obtained	2.9
2.3	TSI Initial and Retest Means, Standard Deviations, Score Ranges, Test-Retest Correlations, and *t*-ratios for Samples of NMR, PS-EMR, and Inst-EMR Pupils	2.10
2.4	Age, IQ, and TSI Means and Score Ranges for 88 NMR, 163 PS-EMR and 108 Inst-EMR Pupils	2.13
2.5	Distribution of TSI Scores in Terms of Frequency and Percent for 88 NMR, 163 PS-EMR, and 108 Inst-EMR Subjects	2.15
2.6	Distribution of TSI Scores Obtained by 88 NMR, 163 PS-EMR, and 108 Inst-EMR Pupils According to Age Categories	2.19
2.7	Distribution of TSI Scores Obtained by 41 NMR, 163 PS-EMR and 107 Inst-EMR Pupils According to IQ Categories	2.20
2.8	Correlation Coefficients of TSI Scores with Initial Teacher Ratings on Nine Scales for 84 Girls and 101 Boys	2.23
2.9	Correlation Coefficients of Initial TSI Scores with Mean Initial Ratings by Six Outside Classroom Raters on Nine Scales for 20 Inst-EMR Girls and 14 Inst-EMR Boys	2.24
2.10	Correlations between Scores on TSI, FTOR, BSRI, and Composite Behavior Ratings on Four T-BRS and O-BRS Scales for 56 EMR Pupils	2.27
3.1	Average Teacher Use of Scale Range in Rating Individual Pupils on Nine Behavioral Rating Scales	3.6
3.2	Interscale Correlations Based on Pretraining Teacher Ratings of 101 Male and 84 Female EMR Pupils on Nine T-BR Scales	3.10
3.3	Correlations between Pre- and Posttreatment Period Teacher Behavior Ratings on Nine Scales	3.13
3.4	Mean Percentage of Rating Scale Intervals Used by Teachers of 15 Classes	3.16
3.5	Correlations between Preratings and Change in T-BR Ratings	3.17
3.6	Correlations between Mean Ratings by Two Sets of Three "Outside" Raters Using O-BR Scales	3.18
3.7	Correlations between Teacher Ratings and the Mean Ratings by Six Outside Raters	3.19
3.8	Interscale Correlations Based on Pre-Treatment Outside Rater Ratings of 34 Inst-EMR Pupils on Nine O-BR Scales	3.20
3.9	Correlations between Pre- and Postratings by Outside Raters on O-BR Scales	3.21

Table of Tables (continued)

Table		Page
3.10	A Comparison of the Correlations between O-BR Mean Preratings and Mean Change in Ratings with the Median of the T-BR Correlations between Preratings and Change	3.21
5.1	Means and Ranges of IQ's and Ages for Trial 1 Junior High School Retardate Experimental, Placebo, and Control (No Treatment) Classes, Singly and Combined	5.3
5.2	Means and Standard Deviations for TSI Pretrial and Gain Scores for Trial 1 Experimental, Placebo, and Control (No Treatment) Classes Singly and Combined	5.5
5.3	Pretreatment Behavior Rating Means and Mean Gains on Five Rating Scales for Trial 1, Experimental, Placebo, and Control Classes, Separately and Combined	5.6
5.4	Average Weekly Ratings by Two Teachers Using Experimental Lessons and Two Teachers Using Placebo Audio-Visual Materials in Junior High School Special Education Classes	5.8
5.5	Mean IQ's and Ages for the Trial 2 Experimental and Contrast Classes, Singly and Combined	5.10
5.6	Pretrial and Gain Scores from the Two Curricular Sessions for the Trial 2, Means and Standard Deviations for the TSI Experimental and Contrast Classes, Singly and Combined	5.12
5.7	Analysis of Variance of TSI Gain Scores for the Two Curricular Sessions for Trial 2	5.15
5.8	Pretreatment Behavior Rating Means and Mean Gains on Rating Scales for Trial 2 Classes Who Used Experimental and Contrast Materials in Opposite Sequence.	5.16
5.9	Means and Ranges of IQ's and Ages for Girls and Boys in the Trial 3 Experimental (J) and Contrast (K) Classes	5.19
5.10	Means and Standard Deviations for the TSI Pretrial and Gain Scores for Girls and Boys in Trial 3 Experimental (J) and Contrast (K) Classes	5.21
5.11	Pretreatment Teacher Behavior Rating Means and Mean Changes on Three Rating Scales for Girls and Boys in Experimental and Contrast Classes	5.22
5.12	Pretreatment Mean Ratings and Mean Changes in Ratings Made by Six Outside Raters of Social Behavior of Boys and Girls in Trial 3 Experimental and Contrast Classes	5.23
5.13	Mean IQ's and Ages for the Trial 4 Experimental Classes, Singly and Combined, and for the Trial 1 Control Class	5.25
5.14	Means and Standard Deviations for the TSI Pretrial and Gain Scores for Girls and Boys in Four Trial 4 Experimental Classes and for a Control Class	5.27
5.15	Pretreatment Behavior Rating Means and Mean Change in Rating for Trial 4 Experimental and Control (No Treatment) Classes	5.29

Table of Tables (continued)

Table Page

5.16 Summary of Significant Differences between Groups in Four Trials in Terms of Changes in T-BR Ratings 5.34
6.1 Means and Ranges of TSI Scores of 163 PS-EMR and 107 Inst-EMR Subjects in Relation to Residence, Age and IQ 6.4
6.2 Means and Ranges of TSI Scores, IQ's, and Ages of 12 PS-EMR Classes and 9 Inst-EMR Groups 6.5
6.3 Mean TSI Score Gains and Ranges in Relation to Sex, Age, and IQ of 125 Experimental Class Pupils 6.8
7.1 Summary of Trials of Social Perceptual Curricula 7.5

TABLE OF FIGURES

Figure		Page
2.1	Picture #28 - Test of Social Inference	2.5
2.2	Distribution of TSI Scores in Terms of Percentage Frequencies for 88 NMR, 163 PS-EMR, and 108 Inst-EMR Subjects	2.16
2.3	Picture #25 - Test of Social Inference	2.17
2.4	Mean TSI Scores for IQ Intervals for 8 NMR, 163 PS-EMR, and 10 Inst-EMR Subjects	2.21
2.5	Pearson Product Moment Correlations for 84 EMR Females and 101 EMR Males between TSI Scores and T-BR Ratings on Nine Scales and Between IQ's and T-BR Ratings	2.25
5.1	TSI Gains by Pupils Using Experimental, Placebo, Control and Contrast Programs, Gains Shown Separately for Boys and Girls	5.32

I. INTRODUCTION

Background

A major rehabilitation research question concerns the characteristics or factors most relevant to the post-academic or post-institutional success or failure of retarded individuals. A number of early studies cite the poor predictive value of the IQ for vocational success. As an example, Kennedy's follow-up of 256 former institutional subjects reported them "equal to a matched group of normals in economic self-sufficiency." (Taylor, 1964, p.3) O'Connor (1953), in England, examined the employment record of 47 adolescents, mean IQ 70, range 65-79, in the two years prior to Army Service and found it comparable to that of a control group, mean IQ 100, range 94-106; although, to be sure, the latter had obtained more skilled jobs. Masland, Sarason & Gladwin (1958) cited additional studies of this nature and from them concluded:

> What is needed is a battery of scales.../to/ predict to some degree... the ability to develop social and occupational skills adequate for social living.../to/ be divorced as far as possible from the IQ concept. (p.306)

Windle (1962), after careful review and evaluation of then existing prognostic studies, reported that a higher Full Scale IQ, and that a Performance IQ higher than a Verbal IQ, was prognostic for institutional release but that neither Performance nor Verbal IQ was sufficiently prognostic of success thereafter. He concluded that research was needed in the areas of patterns of personality or behavioral abilities, to enable a shift from medical to sociological diagnostic categories.

Subsequently, an IQ-success relationship was noted in several studies. In one of them (Jackson & Butler, 1963) Verbal IQ, age, compliant behavior and other attributes were related to successful community placement of former institutional retardates. In others (Deno, et al, 1965; Taylor, 1964) Performance IQ was found related to post school or vocational adjustment. Taylor reported that nine successful vocational groups had higher WISC Performance IQ's than the less successful, and that a "perceptual cognitive factor" he termed "lack of distractibility" underlay supervisor evaluations of both vocational and personality adequacy. (p.77) Although age, sex (Deno, et al, 1965)[1], manual dexterity, and absence of physical handicaps (Taylor, 1964; Guralnick, 1963)[2] have been found to be significant variables, many investigators have noted that personality, or social competency, not well encompassed by existing methods of assessment, is the most crucial factor.

---

[1] Windle (1962) found sex nonsignificant with respect to prognosis, but cited wide differences between male and female adult roles, requiring careful scrutiny of criteria of adjustment.

[2] This differs from Windle's finding that physical disability was not related to post-institutional success or lack of success.

Heber (1961, 1962) proposed a concept of "impaired adaptive ability" as the critical disability in instances of failure. He described adaptive behavior as consisting of the many specific behaviors by which one copes with or adapts himself to the natural and social demands of the environment. He regarded the Vineland Social Maturity Scale as "probably the best single measure." Leland and his co-workers (Leland, 1964; Leland, Nihira, Foster, Shellhaas & Kagin, 1966) in refinement of the adaptive behavior concept are currently developing scales for its measurement. Shellhaas described their approach as one in which the reports of employers and other acquaintances of retarded individuals are used as sources of information as to community behavioral expectancies and tolerances. From such information they have gleaned a number of criterial behaviors as components of successful independent functioning, of occupational adaptation, of social responsibility, and the like. From their viewpoint:

> ...the degree to which /the individual/ satisfactorily meets the culturally imposed demands of the environment determines how well he is... function/ing/. His ability to function is...socially defined by the community in which he resides, works, or participates. (Shellhaas, 1965, p.1)

This concept of adaptation relates to Leland's and Goldberg's (1957) formulation of the rehabilitation process:

> Rehabilitation is a continuous process which utilizes all services working as a team to help an individual with physical, mental, and social disabilities and handicaps to raise the level of his physical, mental, and social efficiency. (p. 530)

Project Objectives

This current project has been based on a somewhat similar view of the rehabilitative goal; that a successful person is one who is able to move in and out of a number of settings and to function acceptably while within them. Through acceptable self presentation and performance in job interviews and on the job, he earns the money by which to sustain himself. Through adequate functioning in public eating places, hotels, boarding homes, stores, etc. he makes himself accepted. Through acceptable behavior in places of recreation and with his friends he elicits the off-the-job satisfactions that are important to his emotional well-being.

The project was generated from the assumption that the retardate's failure to match his behavior to the situation norms (encoding) would be closely related to the slowness, confusion, and inaccuracy with which he interpreted (decoded) social situations in pictorial form. His ability to detect and interpret social cues or signals was viewed as modifiable through training. (Edmonson, de Jung & Leland, 1965).

The project objectives, based on the above considerations, were:

(1) To devise an objective test of social situation comprehension. This test, in connection with the project goal of experimental social perceptual training, would be used as a measure of change associated with treatment. In its more general use the test would aid in the assessment of the cue interpretation deficit of retarded (or non-

retarded) subjects as an indication of their need for remedial training in this area. A further use, subject to investigatory trial beyond the scope of this project, was the possibly predictive relationship between social cue interpretation and the post-training social adequacy of a rehabilitation trainee.

(2) To devise a remedial program focused upon social cue recognition and interpretation, termed interchangeably as social perceptual or as social inference training, and to assess the effects of its experimental use at the prevocational level. The training was conceived of as analagous to methods of familiarizing pupils with geographically distant places. In this instance, in lieu of the geographically distant, it was planned to use materials illustrative of the adult milieux and to discuss and illustrate relevant response behaviors.

(3) To try out the social perceptual training materials in public school EMR and institutional classrooms to determine their effectiveness in improving social situation comprehension and in modifying social behavior. In addition, the materials would be evaluated as to their appropriateness in the total prevocational curriculum.

Related Research

Sociological studies have shown that in social settings, particular patterns of behavior consistently occur and become culturally expected. These patterns, as elaborated by Goffman (1959, 1961, 1963), Hall (1959), Brown (1965), and others, are remarkably precise and narrow with respect to self presentation; in particular, to approaches and transactions with others. There are behavioral requisities; first; for admission, and then for degrees of interactional acceptance. The lack of these social requisities limits admission to and acceptance in social settings. Deficit in this regard, whether for reasons of learning disability or limited situational exposure, is surely as handicapping as academic failure.

In pilot studies it had been noted that the interpretations by retarded adolescents of pictures of social situations differed notably from interpretations by nonretarded subjects of the same age. Although the retarded might respond with literal accuracy, subjects would frequently fail to grasp the full meaning of what they saw. Unlike nonretarded subjects of similar ages, who would restrict themselves to the socially relevant and get to "the point" of the matter, the retarded tended to include the irrelevant, and although they might grasp the implications of portions of a scene, they would typically stop short of the implied relationships.

Picture material used to elicit these differences were scenes of persons performing some action. The position of a depicted person in relation to another, his posture, gestures, facial expression, or dress, were cues as to intentions and feelings. The retarded, apparently less aware of the implications of such cues, tended to describe what they saw in an enumerative way.

> There's a boy and he's looking under the bed. There's a calendar on the wall. There's a man and he's looking at his watch. There's a dog and he's holding a shoe in his mouth.

Sometimes they might approach the point:

> It might be his dad. He might be waiting for the boy. He wonders if it's late.

Sometimes they might make an inference or two and still fail to put them together into the expected relationship:

> The dog's holding the boy's shoe. The man is waiting for something.

The presumed relevance of this perceptual cognitive (cue interpretation) deficit to social behavior may be made clearer by an example. Typically, in entering a social situation, a nonretarded person makes a series of categorical inferences which serve him as guides to behavior. As an example: "A *very busy man* (an inference from voice, expression and behavior) *somewhat upset* (a further inference from voice and expression) *by the conversation* he is having on the telephone" (a further inference as to the cause of the emotion). Consequent to these three inferential discriminations, the nonretarded modulates his behavior. The retarded, however, who sees only "A man--talking on a telephone--and he is sitting at a desk and there are lots of papers on it" is disadvantaged. There is nothing in this interpretation that is prescriptive of a successful approach. This may be a reason for the retardate's social rejection. His acultural behaviors--his approaches and transactional behaviors which are unresponsive to mood and status and to the occasion mark him as an undesirable.

References to the failure of retarded subjects to get intended meanings from pictures are found in the literature describing their performance on a variety of tests. "Stereotypy" (Beck, 1952), "abnormal concreteness," "description" (Beier, Gorlow & Stacy, 1951; Sarason, 1943; Denhof & Robinault, 1960; Butler, 1961), "bare descriptive phrases" (Gorlow, Butler & Guthrie, 1963), are terms that have been used to characterize the inadequacy of their responses. In a global sense the cited inadequacy of the retardate appeared to be his failure to go beyond what he perceives to what the material implies. Bruner (1956; 1958) in describing this process of "going beyond the given" bases his discussion on the assumption that inference begins with the assigning of data to a category. He indicates inference as commencing with the identification, on the basis of certain criterial cues, of an event or object as the number of a class.

> ...we stimulate an /*individual*/ with some appropriate input and he responds by referring the input to some class of things or events. "That's an orange," he states. On the basis of certain defining or criterial attributes in the input...usually called cues although they should be called clues, there is a selective placing of the input in one category of identity rather than another...(Bruner, 1958, 680ff)

Without classification, each situation or item would be regarded as totally new. However, because of ability to identify or infer the similar on the basis of the criterial cues, we can bring to a situation what we have learned at another time in another situation.

The apparent difficulty with which the retarded acquires classifications has been noted in the literature. Silverstein & Mohan (1963) refer to their poor performance on sorting tasks. Stephens, in two papers (1964; 1966) examined the categorization performance of a group of pre-adolescent, non-institutional retarded boys relative to the performance of same aged boys of average intelligence. He reported the retarded as being significantly less able to use expected common categories in grouping pictures of objects and people than the nonretarded comparison group. He concluded that if use of categories is an important tool in attaining meaning from experience, the life experiences of the retarded must be narrow and much of their experience incomprehensible. (1966, p. 361)

Clark & Thomson (1963) demonstrated that a retarded population had an especially poor grasp of categories of social situations. In their study they compared the responses of an institutional population of 127 subjects, ranging in age from 5 to 60 with mean IQ of 60, to two series of 16 pictures each. The series of *social situations* pictures consisted of four each from the categories of crime, danger, play and work. The *object* series consisted of four each from the categories of transport, furniture, plants and animals. Thirty-one, or 24 percent, of the subjects were unable to sort into classifications at all. Ninety-six, or 76 percent, of the subjects performed at varying levels of success with the *object* pictures. In contrast, only 55, or 43 percent, of the subjects could perform at varying, but lesser, levels of success with the *social situation* pictures. No sex differences were found. Correlation between WAIS IQ and successful classification of *object*, although not *social*, was significant at the .01 level. Although the youngest and the dullest of the subjects could not classify either the *object* or the *social* materials, the authors noted that age was a nonsignificant factor in their population.

Studies can be cited which report beneficial results as a consequence of direct training in the use of concepts. Chansky and Taylor (1964) investigated the effect of training retarded pupils to reproduce block designs correctly. In this study of the effect of direct training in veridical perception, the WISC, Benton Visual Retention Test (BVT) and California Achievement Test (CAT) were administered to two experimental classes of 13 pupils each, aged 8 to 11, IQ's approximately 50 to 75, and to a similar control class. Ten weekly sessions were held in the use of a set of Blocksville blocks to reproduce designs. Training sessions of from 45 minutes to an hour's duration were conducted on an individual basis for pupils in one experimental class. The control class had no special training. All groups were retested on all instruments twelve weeks following the initial test.

In addition to the reduction of error with respect to the block design reproductions by pupils in the experimental classes, significant improvements were noted on Similarities and Vocabulary subtests of the WISC as well as on the Verbal Scale, Performance Scale and Full Scale IQ's. Experimental groups made significantly fewer errors on retest with the BVT. The superior improvement of the individually trained pupils over the group trained group was apparent on Blcok Design, Picture Arrangement, and Object Assembly subtests of the WISC.

Chansky notes that these findings agreed in part with his earlier (1963) study with normal children. The authors theorized that

Guiding retarded children to the correct perceptual responses and then

> reinforcing them alters the manner in which they decode stimuli...As they become better organizers and discriminators, they acquire more accurate information...(p.407)

Levi (1965) reported remedial "treatment of a disorder of perception and concept formation in a case of school failure." Psychological testing of an 11 year old boy failing in school work revealed particularly poor visual motor and conceptual performance. Treatment goals in a psychiatric clinic were to teach a store of categories and improved ability to select categories appropriately. Retesting after a year's treatment and a follow-up six months later showed satisfactory school performance.

Bruner's emphasis on the importance of classification as preliminary to being able to think of the implications of the immediately given, together with studies showing classification as one of the deficits of the retarded, would suggest a Montessorian kind of rehabilitative effort in which tasks of sequenced difficulty, grouped in terms of categorical relationships, would be presented. However valuable such approaches may be to academic performance, it does not appear to the present authors that the development of proficiency in discriminating between object categories and in putting objects in relational order would generalize to knowing "what sorts of things to do in certain kinds of situations." Vernon (1955, p.181) in writing of this latter ability, speaks of our gradual acquisition of "schemata," used as guides to perception and action. He uses the term to stand for systems of beliefs and expectancies; for impressions of relationships and regularities in the world. In discussing the role of education in the acquisition of schemata, examples were cited of the value of telling or showing an observer what to look for and then providing practice--the method apparently successfully followed by Chansky and Taylor in their study summarized above.

A study by Hess and Shipman (1965) is more explicit as to the manner in which social behavioral schemata appear to be acquired. The authors examined the behaviors of families, viewed as systems of communication and control. In some families, and especially among those of lower income and lower education, communication to children more typically took the form of peremptory behavioral directives without explanation of the situational context for which the behavior was demanded; as examples: "Be quiet"..."Shut up"..."Mind the teacher." Conformity would be enforced by punishment, less frequently by reward, even less frequently by praise. In some families, and especially among those of higher income and higher education, communication of a behavioral directive would be accompanied by an explanation. "Would you keep quiet a minute; I want to talk on the phone." The difference between these modes of communication was commented upon as follows:

> In one instance the child...is asked...to comply; he is not called upon to reflect or to make mental discriminations. In the other example the child is required to follow two or three ideas. He is asked to regulate his behavior to a time dimension; he must think of his behavior in relation to its effect upon another person. (p.872)

The authors implicitly make the point that information conveyed in this latter manner to children becomes a major source of socially functional schemata by which one is able to regulate one's own behavior, and that in the absence of such schemata either passive conformity or asocial behavior, impulsive and resistive, are frequently observed.

In generalizing to the cognitive and behavioral deficits of the retarded, the Hess and Shipman study suggests the rehabilitative approach of attempting to convey by language, by illustration and by demonstration, veridical concepts of culturally expected behaviors in relation to persons, activities, and settings.

An experimental study reported by Clark (1967)was apparently successful in conveying an adult evaluative frame of reference to a group of institutionalized retarded. The investigators were concerned with the lack of judgment exemplified in retardate behavior:

> ...it seemed...that the poor social adjustment of many subnormals was due not so much to bad judgment as to the fact that no judgment was really made at all, their actions following only on impulse rather than on assessment.[3] (p.31)

Clark's paper described an experimental effort to communicate to the institutional population an awareness of those behaviors which are culturally valued. The procedures resulted in changes in judgments made by retarded subjects in the direction of greater accord with judgments made by members of the staff. The subjects in the study were 70 adolescent boys, ages 13 to 21, mean CA 17. The IQ's ranged from 30 to 90 with mean IQ of 59. Vineland Social Maturity Scales SQ's ranged from 15 to 115 with a mean of 72.

In the study each subject was rated weekly with respect to social competency by five members of the staff. At the end of a baseline period, without informing the residents of the existence of staff ratings, nor apparently, the basis for their selection, the four most highly rated boys were designated as prefects. The prefects, for a period of 16 weeks, were to head up a cottage team of 16 boys, and had the responsibility for discipline and work productivity of their respective groups. The prefects, taking turns, were allowed to choose team members. The order of their selections was recorded and used as the basis for the comparison of the "peer judgments" with the staff ratings. During four successive 16 week intervals the weekly staff ratings of social competencies were continued. In addition they were posted in the cottages where they would be visible to the residents and the reasons behind the ratings were explained. At the end of each 16 week interval the boys who received the highest staff ratings were designated as prefects, who then chose their teams for the next period. The rank order of their selections on each occasion was recorded for comparison with staff assessments.

---

[3] Although other studies (de Jung & Haring, 1962) have suggested the greater fluctuation, the lesser reliability, of evaluations made by mentally retarded than those by nonretarded, it is not correct to assume that evaluational schemata are lacking. In the Clark Study itself, although the initial selection of team members was made prior to their knowledge of the staff ratings, the rank order of peer selections, converted to scores, was found to correlate with staff ratings with a coefficient of .60.

The result, in terms of coefficients of correlation between the prefect's selections of team members at the end of each four months and the ratings of residents by members of the staff was an increase in agreement between evaluation. The baseline rank order correlation between staff assessment and peer selection for the period prior to the posting and discussion of staff ratings was .60. The correlations between staff assessments and the peer selections for the subsequent four 16 week periods were .83, .68, .71 and .81 respectively.

Project Overview

The four years of this project consisted of several phases in which emphasis would be placed upon one and then another of a series of subgoal activities. The initial phase involved development of the Test of Social Inference (TSI). Because the development of this measure of social comprehension was one major activity of the project, the account of the procedures that were followed in its formulation, together with a description of its administration, scoring, and discussion of the relationships between TSI scores and other measures, is given separate mention in Chapter II.

Second phase activities were initiated by the need for a method of assessing change in overt behavior which might be associated with the project's experimental social perceptual training. A resume of activities centering around techniques for observations and ratings of behavior is contained in Chapter III. This chapter describes the two rating-scale formats used in all of the experimental trials as pre- and post treatment measures of the effect of the social perceptual training.

Third phase activities focused on the development of a method by which to improve the retardate's social comprehension. In Chapter IV are described the approaches that were considered and the procedures and materials which were developed and given experimental trial by the project.

The fourth and major phase covering a thirty-month period involved the trial of the social perceptual materials with sample groups of retarded adolescents at the prevocational level, followed by revisions of the material. During this phase, sets of illustrated lessons were given trial in junior high school special education classes for the retarded. As a general practice, the use of the experimental curricula was compared with results of use of a contrast program. Trial of each version of the experimental materials was followed by extensive revision, based largely upon the suggestions and criticisms of the participating teachers and project consultants. A description of the trials and their results will be found in Chapter V.

A resume of the overall results in relation to the objectives of the project is contained in Chapter VI, together with a discussion of their implications for rehabilitation programs for the retarded adolescent. Chapter VII gives an overview of the entire project proceedings and results.

II. TEST OF SOCIAL INFERENCE (TSI)

A major project goal was to devise a test of social situation comprehension to compare performance across subjects and by means of which to measure the effects of remedial treatments. The central ingredients of such a test were suggested by earlier studies of the responses of EMR junior high school age children to picture stimuli. As noted in the preceding chapter, retarded subjects in comparison with nonretardates had been found to perform very poorly in their comprehension of pictured social situations. The Test of Social Inference (TSI), introduced in this chapter, was developed to assess the retardate's "social perception" in terms of his interpretation of social cues. The first sections of this chapter describe the extensive project activities relating to item selection, methods of test administration, and scoring. Further sections present data relevant to interexaminer agreement, reliability of scores, and the relationship of the TSI score to other assessments.

A. Early Development

An initial study by the senior author investigated the responses of retarded and nonretarded subjects to a group of pictured social situations. The subjects were six nonretarded and 18 retarded boys and girls ranging in age from 4 to 15. In the examination procedure the examiner displayed 65 pictures, chosen from news media as representative of a wide range of social situations. The pictures were shown, one at a time, with some variant of the request, "Tell me what this is about." There was no time limit for response. Sessions ranged in length from 20 to 50 minutes,and were tape recorded.

This trial testing revealed little difference between the retarded and the nonretarded or between the younger and the older subjects with respect to accuracy of report. The difference was in succinctness, the ease with which the older and the nonretarded subjects could get to the heart of the matter. The younger and the retarded subjects would "roam around" in a picture, apparently considering everything as of equal weight and interest, responding, as an example:

> There's a girl. There's a baby. They're sitting on something. She's holding a piece of bread with butter on it. There's a house in back of them...

At times after naming the details, a summative interpretation of the relationship might be attempted, such as:

> There's a boy; there's a girl; there's a piece of bread; the little baby is trying to get the bread away from the little girl...

At other times no indication would be offered of comprehension of the interpersonal situation.

This exploratory study emphasized that accurate noting of detail is a necessary part of eventual social comprehension, but that additionally required is the recognition of interactional relationships, of intentions, of whatever is implicit in an event that is apt to give it movement and direction and to require response from a participant. Social comprehension requires awareness that some clusters of details have functional implications that are not possessed by others.

There were three criteria for test stimulus pictures which would elicit responses that would differentiate between subjects performing at the "enumeration-detail" level and those performing at the level of implication. The first criterion was that an event be illustrated whose nature or state could be inferred from the depicted relation of person to person, or person to object, via such details as facial expressions, postures, dress, and characteristics of setting. As a second criterion, pictures should lend themselves to consensual agreement as to their meaning. As a third criterion, the pictures should illustrate a wide range of situations.

Pictures meeting these criteria were the fruit of systematic search. With the assistance of the collection department of Goodwill Industries, thousands of copies of popular magazines were secured. Among them were American Home, Boy's Life, Ebony, The Farm Journal, Glamour, Ladies Home Journal, McCall's, Playboy, Presbyterian Life, Scouting, True Romance, and many others. All of these, and virtually complete files of back issues for at least ten years of Life, Look, and Saturday Evening Post were examined. From a pool of some 130 pictures, selected groups of items which required at least two, and preferably three or more, inferences for a full interpretation of the pictured events, were tried out in a series of trials with retarded and nonretarded subjects. At the same time, variations of questions and methods of response recording were tried. Using the responses of the nonretarded adolescents, pictures which proved ambiguous were eliminated, as were pictures reflecting clear sex differences favoring either boys or girls. This pilot testing continued over a period of three months. In all, some 25 nonretarded (NMR), 39 educable mentally retarded (EMR), and 7 trainable level retarded (TMR) between 4 and 25 years of age were tested. A summary description of these pilot trials is presented in Table 2.1

TABLE 2.1

Resume of Pilot Trials for Development of TSI

Pilot Trial	Subjects[a]	Purpose
1	12 NMR, 6 EMR	Comparison of tape recorded responses by EMR and NMR subjects to 65 items
2	9 EMR	Comparison of reliability of tape recorded retest responses to two administrations of same items
3	10 EMR	Comparison test-retest responses by EMR subjects to 41 items
4	3 NMR, 2 EMR	Comparison of checklist recordings with tape recorded responses
5	7 NMR, 10 EMR 7 TMR	Trial of revised check list scoring. Investigation of adequacy of the instrument "floor" and "ceiling"
6	3 NMR, 2 EMR	Trial of revised check list and administrative procedures

[a] NMR - Nonretarded; EMR - Educable Retarded (IQ's 55-85); TMR - Trainable retarded (IQ's 35-55)

The pilot trials revealed the importance of specific questioning as a guide for subject response. Without direct questions, the fact that a subject failed to respond with an expected inference could not be interpreted as inability to make the inference. It could mean that he had not properly construed his response task and had assumed that a descriptive or peripheral answer was all that was required of him. In the pilot testing, "reminding" questions were asked of the subject when each picture was presented, e.g. "Who is she?" "How does she feel?" "Where is he?" "What is he doing?" "Why is he doing it this way?" When the attention of the retarded subject is directed in this prosthetic way through examiner question, he responds with more of the relevant social inferences than he does when the convergent question is not asked him. In the final procedures developed for administering the TSI, presentation of the first half of the pictures was accompanied by specific questions of this kind, but not so for the last half of the pictures.[4] In presenting the latter pictures the subject was given the more general instruction "Look at this carefully and tell me about it."[5]

Initially, because of the linguistic deficiencies of most retarded subjects, the possibility of having subjects indicate their comprehension nonverbally was seriously considered. A reasonable "nonlanguage" procedure would be to ask the subject to select the one from a group of multiple choice pictures which would best represent the central social import of the stimulus picture. Several tests have successfully used this mode of response in testing younger and/or handicapped youngsters. The concern in the present testing was to reveal as clearly as possible the subject's own comprehension or structuring of the depicted situation. This objective is not attained when his task is to make a choice between given alternatives. It is also to be expected that any given set of multiple choice items would introduce confounding bias due to the content of the particular choice items. For these reasons, the open end verbal response appeared preferable for examining inferential performance.

The decision to use verbal responses as representative of a subject's comprehension (decoding) was arrived at after extensive examination of responses. TSI protocols from retarded pupils of junior high school age were carefully scrutinized to determine if it were possible to credit subjects with the right idea, apart from the form in which it was expressed. Although consensus between evaluators was not always attained, a clearer formulation of bases for a decision was achieved. This formulation, together with examples of scorable and nonscorable responses was organized into a reference Scoring Guide. For each test picture, those responses which appeared to denote a particular inference, regardless of the form in which it was expressed, were grouped together. From this comprehensive and unwieldy listing of responses and assigned scores, successive "editions" of abridged guides were developed which more narrowly defined response categories in relation to score, and gave examples of each.

---

[4] The test as used throughout the project was not revised to eliminate several less effective items among the 36 test pictures. Two items early noted as being somewhat less discriminative, were placed at the end of the test in a position advantageous to the possible recomputation of data contingent upon their elimination from the set of pictures. In exception to the administration above described, these two items among the latter group of pictures are presented with specific examiner questions typical of the first half of the test items.

[5] Because the absence of outside assistance is more typical of real life social situations, it is possible that the retardate's response to the last half of the TSI pictures is more closely related to his overt behavior in a social situation than his responses to the first half. Data relevant to this consideration are being explored in a University of Oregon doctoral dissertation on the TSI by Ronald Smith.

B. Response Classifications[6]

The TSI Manual distinguishes subject responses by assignment to one of three major classifications for scoring: (1) scorable inferences (Inf) of which there are two subtypes, (a) good inferences and (b) ambiguous, incomplete, or insufficiently specific inferences; (2) nonscore responses, and (3) errors (Error). TSI picture #28, reproduced as Figure 2.1, provides referents for this scoring classification.

Bruner suggests (1958, 680ff) that inference is in part the process of classification, as when one identifies something through use of criterial attributes. The response to Figure 2.1, "He is a PFC in the Army/ Marines," is an example of a response that contains two classificatory inferences. The scene in the picture includes a man's jacket hanging on the back of a chair. It is a uniform jacket and there is a chevron on one sleeve. In this instance a very small detail yields information about the person which clinches identification and shortcuts (closes out) alternative hypotheses that could be generated by other items in the room. In addition to the classification aspect of inference, Bruner (1956) mentions the further definition of inference as an "inductive leap" (p.14) from cues visible in the picture to what they most reasonably imply. The further response to Figure 2.1, "He's home on leave," illustrates the "leap." The two additional good inferences contained in this statement, "He's home," and "on leave," are supported by the uniform jacket, furlough bag, the room furnishings, and the man's posture and expression. To be scorable, inferences from the test pictures must be supported by visible cues, or at least not contradicted by them, and must be joined together in a way that is probable. Those inferential statements more relevant to overall comprehension of a given test picture are termed "good" inferences and are each given a full score point per inference. The above examples would each earn two points for a total of four points. Inferences which suggest comprehension of centrality, but which are ambiguous, incomplete, or vague, receive only half-point credit. An example with respect to Figure 2.1 would be the response, "He's lying on his bed like he used to." This shows awareness that the man has been away, but one remains in doubt

---

[6] A more complete discussion of response classifications and scoring procedures is provided in the TSI Manual (Edmonson, de Jung, Leland & Leach, 1966). This manual was followed for all TSI administrations described in this report.

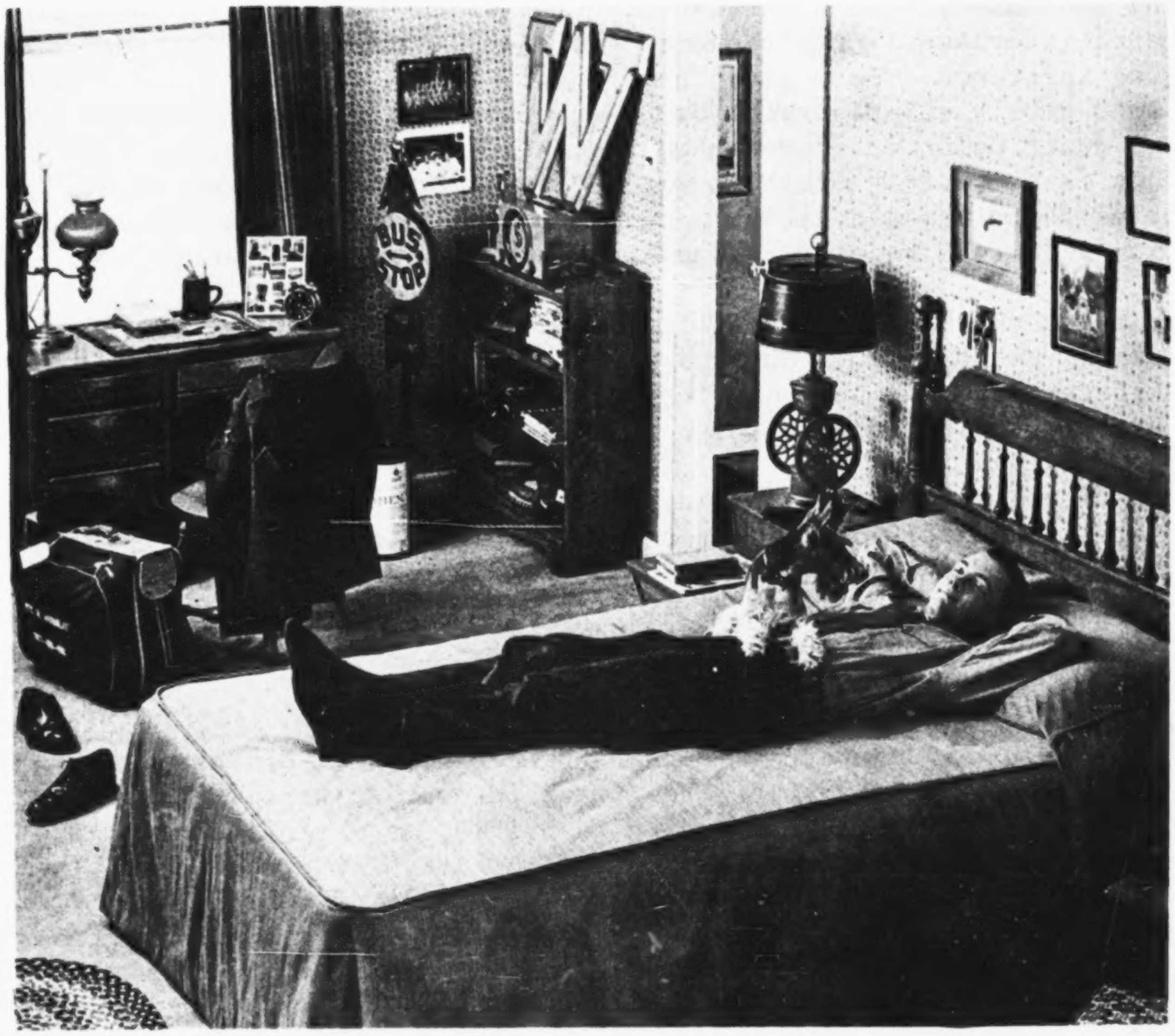


Fig. 2.1 Picture #28 - Test of Social Inference

(Reprinted with permission Simmons Mattress Company
Copyright 1959)

whether the respondent had inferred that the man had been away on military service.

Inference as to irrelevant portions of a scene, low level inferences that fail to differentiate the less well performing subjects from the better, inferences which are of low probability, in addition to descriptive or enumerative responses, are all nonscorable. Examples of irrelevant inferences to Figure 2.1 are the responses, "It's 3:25," or "It's afternoon." These statements are peripheral to the point of the activity of the scene and indicate little understanding of what is going on. An example of a low level inference is the response "He's comfortable", as contrasted with the good inference, "He's glad to be back." A low probability inference, because poorly substantiated by cues, would be, "He's dreaming of his girl." The description or enumeration response involves simply the reporting of what is concretely visible, without implication. The response, "He's lying down; has a dog; has a radio; books, etc..." is an example of description and would receive no credit toward the subject's score.

The third major classification of subject responses to TSI pictures, Errors, although apparently separable into various error types,[7] was not differentiated in tallying the error responses by subjects tested in the course of this project.[8]

Error responses were recorded apart from Inf scores and summed to yield a separate total Error score. An initial intent to subtract Error from Inf score proved unsatisfactory because of the inverse, nonlinear relationship between the two scores.[9] At present writing, except for maintaining records of response behaviors, Error is not being used for measuring subject inferential ability. Numerous references to the TSI Inf score will be made throughout this report. For simplification, this score will be referred to as the TSI score except where both Inf and Error scores are being discussed.

---

[7] Although certain errors appear to be perceptual mistakes, as if a subject's visual acuity or figure-ground perception were at fault, the more frequent errors are those easily identified as misinferences. The responses to TSI picture #28, "He's a bus driver," or "He's in a room in a hotel," are examples of misinference. A study of more refined distinctions among error responses, such as over-generalizations from a particular cue, failure to consider overriding cue information, failure to integrate cue information from several portions of the picture, while of interest to the test developers as means of pinpointing sources of error, has not to date been undertaken.

[8] Bass (1967) found TSI-Error to be moderately correlated (r=.53) with an Inappropriate Behavior Score derived from observed samples of behavior of a relatively homogeneous group of 12 institutionalized male subjects, aged 16-21. Error was significantly inversely correlated with the Appropriate Behavior Score for the same subjects (r=.-71)

C. Test Description - 1967 Revision

In its current form the TSI is a set of 36 large pictures depicting a variety of social events taking place in diverse contexts. The first picture is used as a practice item and is not scored. The test is individually administered.[10] Sessions have run from 20 to 45 minutes. To establish the nature of the task, accompanying the presentation of the first 16 and last two pictures, the examiner asks specific questions. With respect to the initial two pictures, the examiner attempts to focus attention on the relevant cues by asking "What gives you that impression?" Later pictures are accompanied with a more general request to tell what the picture is about.

Responses are recorded in a test booklet which, in addition to listing the questions and probes for each picture, has check lists and boxes for the notation of the three types of responses: Inf, nonscored responses, and Error. The more usual inferences are listed in one column headed "Scorable Answers." The frequent enumerative or other nonscored responses are listed in another column. Errors are listed in a third column. If a subject's response contains any of the listed responses, they are encircled or underlined on the form. Responses that are not listed are written by the examiner in a space provided for the purpose. A sample page of the standard response recording form is included as appendix A. Detailed instructions for the TSI administration are available in pamphlet form. (Edmonson, de Jung, Leland & Leach, 1966).

---

[9] Their relationship is complexly curvilinear. Retarded subjects tested during the test development program (CA's 11 to 20) having Inf score in the range of 86 to 105 typically made six or fewer misinferences (6 or fewer Error points). At the opposite end of the Inf score continuum, some subjects who earned few Inf points made few E points because much of their response consisted of enumeration. Subjects whose Inf scores were in the middle range of 26 to 75 made as few as two or three misinferences to as many as thirty (2 to 30 E points). If the E score were subtracted from the Inf score, the composite total score of the subjects who were actively trying to comprehend what they saw, would, in many cases, fall below the composite scores of subjects who did little more than describe a situation.

[10] A procedure for group administration of the test (in a 26 picture form) using the pictures in the form of slides projected on a screen in the classroom, was given trial with 65 seventh and eighth graders in three classes. Group administration, however, introduced a number of sources of possible error, and the procedure was discontinued.

Upon completion of test administration, the various assignments of one credit per good inference, one-half credit for less complete or ambiguous inferences, and zero credit for irrelevant, improbable inferences and for descriptive responses may be made by use of the printed responses in the test booklet or by use of the Scoring Guide. The subject's total TSI score is the sum of Inf score credits earned in a given test administration. As scored throughout the project, there was no upper limit of acceptable inferences that might be credited for each picture. A few complex pictures elicited as many as seven or eight scorable responses. Although, for this reason no maximum total score could be assigned, an Inf score of 150 was estimated as a theoretical maximum for Inf points.[11] A score of 124, earned by a 14-year-old nonretarded ninth grader, was the highest score actually attained by any pupil tested. An Error score of 50 was estimated as a practical maximum of error points. Although Error is not limited by picture content, it is related to subjects' conceptual schema and verbosity. An Error score of 45, earned by a 26-year-old brain-damaged subject in an institution for persons having neurological impairments, was the highest Error score actually attained by any subject tested.

D. Interscorer Reliability

To determine how successfully the scoring method could be followed by a novice, sets of protocols were scored, first by an inexperienced scorer and, again, by an experienced scorer, both using the Scoring Guide. In the scoring of protocols from the initial testing of ten subjects, the mean of scores assigned by the novice (41.6 TSI points) differed by 2 TSI points from the mean of the scores assigned by the experienced scorer (43.6 TSI points). In scoring the retest protocols for the same subjects, the mean scores assigned by the novice and the experienced scorer of 40.8 and 45.1 TSI points, respectively, differed by 4 TSI points. The sets of scores correlated .99 and .98. Differences between the means were significant, however, using the $\underline{t}$-test for correlated scores, at less than the .05 level of confidence. The former novice, now more experienced, was given a review of her scoring errors and assigned the task of scoring an additional 36 protocols. These protocols were then rescored by the experienced scorer. The correlation between the sets of scores was .99. Given this additional practice, the mean of the scores assigned by the former novice differed by nine tenths of a TSI point from the mean of the scores assigned by the experienced scorer, means of 57.1 and 58.0 for the novice and experienced scorer, respectively. This difference between the scorers did not attain the .05 level. Throughout the project phase in which the test was used as measure of pupil training in social comprehension, test protocols were scored by experienced scorers so as to minimize scoring error. As protection against scorer bias, the scorers were not informed as to the designation of the groups.

---

[11] In a subsequent revision of scoring (Edmonson, Havens & Carrell, 1967), in addition to further shifts in acceptability of certain responses, a maximum score was set for each picture. This establishes a total possible score of 130 points.

E. Test-Retest Reliability

(1) The reliability of subject responses to the TSI was examined in terms of test-retest administrations to several samples of adolescent subjects. These samples were 45 nonretarded (NMR) 8th and 9th graders in a suburban Kansas City public high school, 15 educable retarded (EMR) pupils in an ungraded class at intermediate level in a suburban school in the same area, 45 educable mentally retarded (EMR) from three special education classes in nearby public junior high schools, and 10 retarded patients in a state institution. All tests were individually administered.[12] Ages and IQ's for the various subgroups comprising the four test-retest samples are presented in Table 2.2.

TABLE 2.2

Samples of Non Retarded and Retarded Subjects
On Whom Test-Retest TSI Scores were Obtained

Group	N	Test-Retest Interval	CA Mean	CA Range	IQ Mean	IQ Range
NMR 8th & 9th Gr.	45	1 wk.	180	165-198	NA	
EMR Intermediate	15	1 wk.	172	141-199	69	60-79
EMR Jr.H.S.Class AA	13	9 wk.	200	184-229	64	51-90
EMR Jr.H.S.Class BB	20	9 wk.	205	186-219	66	52-90
EMR Jr.H.S.Class CC	12	9 wk.	178	147-209	70	60-79
EMR Institutional	10	10 days	215	203-243	64	50-84

NA Not available

The NMR pupils in Table 2.2, 22 girls and 23 boys, were retested one week after the administration of their initial test. They were chosen for testing as representing all of the school's 8th and 9th grade "Low English Class" pupils. No intelligence test scores were available for these pupils. The EMR Intermediate pupils, 8 girls and 7 boys, also had a test-retest interval of one week. The EMR institutional subjects, 5 girls and 5 boys, were a few years older than the preceding samples. Their test-retest interval was approximately 10 days. This latter sample was tested by a member of the institutional staff. All other pupils were tested by the project staff.

---

[12] Testing procedures are described in greater detail in the TSI Manual (Edmonson, de Jung, Leland and Leach, 1966) available from the senior author.

Product moment correlation coefficients were computed between the initial and retest TSI scores for each class. These coefficients, together with the means, average gains, standard deviations and ranges of the TSI scores are presented in Table 2.3. With the exception of the retest correlation of .84 for the youngest EMR class, all coefficients for the EMR classes were .90 and above indicating very stable ordering of pupils within these classes with respect to the two TSI scores.

TABLE 2.3

TSI Initial and Retest Means, Standard Deviations, Score Ranges, Test-Retest Correlations, and t Ratios for Samples of NMR, PS-EMR, and Inst-EMR Pupils

		TSI Scores								
		Mean			S.D.		Range			
Sample	N	I	II	Gain	I	II	I	II	r	t
NMR 8th-9th Gr.	45	86.2	93.5	7.3	15.0	14.5	50-116	60-124	.74	4.70*
EMR Intermed.	15	55.2	56.3	1.1	11.1	9.5	38-76	42-76	.84	.73
EMR Jr.H.S.Class AA	13	59.7	63.2	3.5	18.0	10.7	37-94	46-82	.92	1.29
EMR Jr.H.S.Class BB	20	62.6	69.3	6.7	15.9	19.2	36-86	26-104	.90	3.33*
EMR Jr.H.S.Class CC	12	55.2	59.3	4.1	14.7	13.8	27-79	34-90	.91	2.33*
EMR Inst. Class	10	41.8	42.6	.8	19.9	17.1	10-82	12-71	.97	.59

* Significant at <.05 level

For the NMR group, the retest coefficient, an r of .74, was considerably lower. The possibility that this lower correlation coefficient might be characteristic of pupils initially scoring in the higher TSI score ranges is not as clear as the Table 2.3 data might suggest because of the conditions under which pupils of this sample were tested. The retesting, approximately a week after the initial testing, was performed by a second set of examiners and was accomplished in competition with academic examinations and end of school activities. A second factor which may have adversely affected test score stability was the use of novice examiners. Three of the five examiners used to test and retest the 45 NMR pupils had previously worked with the TSI in its development; two were new to the test except for a half-day training session. Since the examiner-pupil test assignments were deliberately arranged for the NMR retests so that no pupil was twice tested by the same examiner, in many instances pupils were tested both by an experienced and an inexperienced examiner. Although the testing schedules for the various EMR classes also employed several (three) examiners and followed a generally similar alternation of test and retest examiners. all of the EMR examiners had received more extensive practice and supervision in use of the test. The possibility of a relationship between TSI response variability and high initial test score cannot be separated without further examination of NMR TSI performance from the possibly operant factors of pupil motivation and examiner error.

As can be noted from Table 2.3, the initial means for the EMR classes ranged from the low 40's to the low 60's; the corresponding mean for the NMR pupils was 86.2, nearly two standard deviations above that of the average EMR class. Examination of individual scores further supported this TSI score distinction between NMR and EMR pupils: Only one of the 45 NMR pupils had an initial score below 58.6, the combined mean of the 60 EMR pupils. Similarly, only one of the 60 EMR pupils in these classes achieved a score above the 86.2 average score of the NMR group.

(2) Gain: Although Table 2.3 presents very satisfactory retest correlation coefficients for all EMR samples tested, both for the one week and nine week test interval, the fact of gain (higher retest scores than initial test scores) must also be considered in describing reliability of measures. For all six of the "reliability" samples higher TSI means were obtained upon retest. As noted by the asterisked values in the rightmost column of the Table, three of these mean gains were significant at the .05 level. Inspection of individual scores revealed no clear association of higher gains with pupil sex nor with pupils whose initial TSI scores were relatively low or relatively high. Although this latter observation casts some doubt upon the hypothesis that the lesser reliability of the NMR pupil performance might have been a function of higher level performance,[13] the possibility is not closed out because of the significantly higher average performance of the NMR pupils. In other words, the high scorers among the EMR perform in a lower range which presumably means lower cognitive functioning, than is true of the high scorers among the NMR. The conclusion of predominantly higher average TSI scores on the second testing holds for all samples; only 21 of the 60 EMR pupils and 6 of the NMR pupils had lower retest scores as contrasted with 45 EMR pupils and 38 NMR pupils who had higher second scores. Although possible general (non-pupil initiated) changes in the direction of more lenient, supportive examination practices and scoring on the part of the several different examiners cannot be fully discounted, the fact that all subsequent testing of EMR samples (See Chapter V data) yielded similar group retest gains supports the conclusion of improved test performance as a function of repeated testing. That the unevenness of this increase was greater for the NMR sample has already been discussed. Whether retest TSI gains would continue with repeated retesting is conjectural at this writing. A possible stabilizing procedure with respect to retest gain would be to introduce the testing procedure with several practice TSI items instead of one. In overview, however, the fact of the retest gain is not a serious handicap to interpretation and utility of the TSI. The high test-retest correlations obtained for EMR samples indicates a reliable across time discrimination among subjects scoring at differing TSI performance levels, for

---

[13] As, for instance, a response pattern of inconsistent, exploratory, hypothesis testing in non-certainty situations.

periods at least as long as the intervals of one week to two months involved in this phase of project investigation. The problem of relating change in TSI to training or other time spanning variables simply necessitates the use of testing control groups to assess the TSI retest changes independent of treatment, a recommended procedure in any case.

F. TSI Validity: Sampling, Age, and IQ Differences

As Ebel (1961) has pointed out, the problem of determining the validity of a test frequently reduces to an unresolvable and pointless questioning of correlation with nonexistent or limitedly acceptable criterion measures. This is in large part true for the Test of Social Inference. The development of the TSI proceeded from the repeatably demonstrated datum that retardates frequently miscue a social interpersonal situation; that they behave inappropriately to the degree that they are less well liked by their peers (Johnson, 1950; Baldwin, 1958), are viewed as poorly adjusted in social situations (Wunsch, 1951; Wallin, 1956; Garfield, 1963), have on-job difficulties (Gellman, 1957; Peterson & Smith, 1960) and, of course, may be so generally unacceptable to the nonretarded community as to be institutionalized (Guskin, 1963). This failing is not true to the same extent for all retardates and, of course, carries over to varying percentages of nonretarded persons. The goal in developing the TSI was to devise an objective test administrable to the retarded adolescent which would reliably identify persons in terms of their ability to make appropriate interpretations of social cues. The satisfactorily high test-retest correlation coefficients reported in the previous section indicate that the TSI identifications are stable; persons ranking lower or higher within their test group with respect to their TSI score tend to remain in that ranked position. In its traditional sense, the further question of validity of the TSI requires criteria measures of social cue interpretation, the lack of which led to the development of the TSI in the first place.

The more meaningful, certainly more answerable, questions at this juncture revolve about the relationship of the TSI score with other subject description. What, in effect, can be said about subjects scoring high or low on the test measure; how do they test out or behave in environments other than that of the particular testing situation? The three questions to be asked regarding the TSI score in the remainder of this chapter are of this nature: (1) The first question concerns differences between the three populations tested, educable mentally retarded pupils in public school special classes (PS-EMR), institutionalized retarded of approximately the same IQ range (Inst-EMS) and nonretarded public school youngsters of approximately the same chronological age (NMR). (2) The second question concerns (a) the relationship of TSI score and pupil age and (b) the relationship of TSI score and IQ. (3) The third question concerns the relationship of TSI scores and behavior ratings made by teachers or other official observers. (4) The fourth question concerns the relationship of TSI scores and scores earned on the Vineland Social Maturity Scale, the Fudell Test of Occupational Readiness and the Bown Self Report Inventory. Data relevant to these four questions are presented in this section.

1. TSI Scores in Relation to Sample Differences: An underlying assumption of the project in developing a curriculum to improve the social inferential performance of the retarded is that his performance is modifiable; that attention

discrimination, decoding, and appropriate and inappropriate responses to social cues are learned. A corollary assumption involves the past and present experiential history of the subject; i.e., that his experiences and their outcomes (aversive or satisfying) largely determine the range of his social situation functioning. Mean differences in social cue inferential performance might be anticipated between the NMR, PS-EMR and Inst-EMR because of differences both in the range and outcomes of the social experiences available to persons in these different groupings. Although no inventory of social opportunities available to NMR and EMR pupils can be cited, Barker and Gump in their social ecological studies of school pupils observed that "marginal" pupils (below 90 IQ) participated "in fewer school behavior settings." (1964, Ch.8). Although the differences in participation between pupils of lower and higher IQ were most extreme in large schools, it was found that regardless of school size, pupils having IQ's of 90 and below participated in many fewer social situations than pupils whose IQ's were 100 to 110 and above (Ibid. p.88). These authors list and discuss reasons given by participators for their participation (Ibid. p.197). Omitted is similar reference to reasons for nonparticipators' nonparticipation. It requires little imagination, however, to assume that the consequences of participation by the less knowledgeable participator differ from the consequences of participation by the more sophisticated participator and that these differing contingencies work to increase or decrease participation.

No data concerning the daily experiences of retarded subjects in institutions is available to the authors. It seems a reasonable assumption, however, that the behavioral deviance more frequently noted among institutionalized subjects would have limited their social exposure prior to their placement and would have created many adverse outcomes to their experiences.

The expected difference between the three groups in social cue interpretation is supported in terms of TSI scores obtained from the populations tested, as is revealed in Table 2.4.

TABLE 2.4

Age, IQ, and TSI Means and Score Ranges for 88 NMR, 163 PS-EMR and 108 Inst-EMR Pupils

		Sex		CA		IQ		TSI Score	
Sample Group		M	F	Mean	Range	Mean	Range	Mean	Range
NMR 8th & 9th	3 classes	49	39	177	156-198	102[a]	82-119[a]	83*	53-116
PS-EMR	12 classes	99	64	178	141-229	69	45-90	55*	8-104
Inst-EMR	9 classes	65	43	199	113-244	64[b]	43-87	39*	5-84

---

* Differences between means $<.05$ l.c.

[a] IQ based on 41 subjects in two classes; not available for remaining subjects.

[b] IQ based on 107 subjects. IQ not available for one subject.

The data in Table 2.4 are summaries of the ages, IQ's and TSI scores for the initial, or, in some instances, single time testing of 24 groups and classes of adolescents as part of various project studies. In all, the data for 359 pupils are summarized in Table 2.4. The 88 NMR pupils include the 45 eighth and ninth-grader "low English class" pupils previously mentioned in Section E of this chapter. The additional 43 eighth and ninth-graders received a single administration of the test. These additional pupils were in a lower social economic area in a Kansas City, Missouri public high school. The 13 ninth graders, all boys, all low achievers, were undergoing an experimental program as "potential school dropouts." The remainder of the pupils were low achieving eighth graders.[14] The 163 PS-EMR pupils were from 12 special education classes; 10 of them in the Kansas City area, and 2 of them in the Springfield, Missouri public schools. The 108 Inst-EMR pupils were from four special education classes, and three smaller specially constituted experimental groupings at the Parsons State Hospital and Training Center, Parsons, Kansas, and from two classes at the Hissom Memorial Center, Sand Springs, Oklahoma. The individual class distributions of age and IQ were reasonably similar within each of the three major groupings in Table 2.4. The summary means and score distributions for the three major groupings are presented in the table.

As may be noted from Table 2.4, the expectation that these three groups would differ with respect to their TSI performance is borne out by TSI means of 83, 55, and 39 for the NMR, PS-EMR, and Inst-EMR groups, respectively. Critical ratios for the differences between the means were all well above that required for significance at the .05 level. That these differences were obtained in spite of the slightly higher average age of the Inst-EMR sample, even more strongly asserts the conclusion of intergroup differences.

The extent of group separation and overlap with respect to individual TSI scores is shown by frequency distribution for the same subjects as presented in Table 2.5.

---

[14] These 43 pupils had been identified two years earlier in a research study as potential dropouts on the basis of four criteria: low school achievement, Otis IQ's under 100 and teacher and peer nominations of aggressive maladjustment. The experimental ninth grade treatment had included a supportive academic program, counselling, and extra-curricular involvement with a group of adult males. The eighth graders were receiving regular school program.

TABLE 2.5

Distribution of TSI Scores in Terms of Frequency and
Percent for 88 NMR, 163 PS-EMR and 108 Inst-EMR Subjects

TSI Score	NMR f	NMR %	PS-EMR f	PS-EMR %	Inst-EMR f	Inst-EMR %
110-119	2	2	0	0	0	0
100-109	7	8	1	1	0	0
90-99	22	25	3	2	0	0
80-89	19	22 ◁	6	4	3	3
70-79	19	22	30	19	6	6
60-69	16	18	25	15	8	7
50-59	3	3	46	28 ◁	8	7
40-49	0	0	20	12	23	21
30-39	0	0	20	12	25	23 ◁
20-29	0	0	7	4	19	18
10-19	0	0	3	2	11	10
0-9	0	0	2	1	5	5
Total	88	100	163	100	108	100

Between the NMR and the EMR samples the distinction is clear in terms of the generally lower TSI performance of the retarded. Of the NMR sample only three percent scored as low as the mean TSI score of the PS-EMR pupils and only six percent of the latter scored as high as the NMR mean. Of the PS-EMR sample some 20 percent scored as low as the Inst-EMR mean. Of the Inst-EMR sample approximately the same percent scored as high as the PS-EMR mean. Only three percent of the Inst-EMR pupils scored as high as the NMR mean.

A graphic presentation of this separation and overlap of scores is provided by Figure 2.2 in which the percentage frequences from Table 2.5 are plotted.

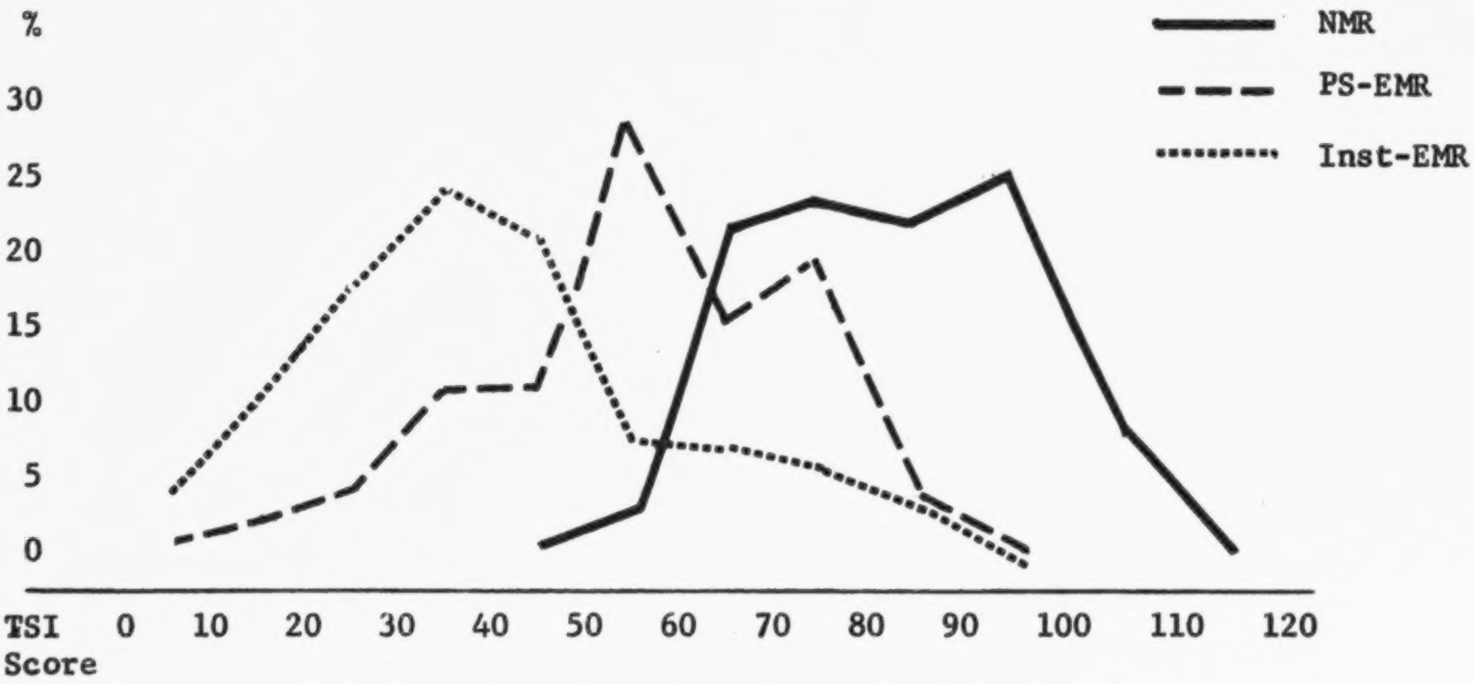


Fig. 2.2 Distributions of TSI Scores in Terms of Percentage Frequencies for 88 NMR, 163 PS-EMR, and 108 Inst-EMR Subjects

From the overlap of the distributions in Figure 2.2 it can be seen that a number of retarded adolescents in special education classes and institutions were indistinguishable from same aged NMR youths[15] in their ability to interpret social cues. This is interpreted as supporting the project's contention that although TSI performance is usually lower among retarded subjects, there is as yet no evidence that factors responsible for retardation prevent the acquisition of social cue interpretation if appropriate learning experiences are provided.

In marked contrast to the "normalcy" of the performance of the high TSI scorers, at the other extreme of the range of EMR scores, however, the test protocols suggest how little awareness these persons have of what transpires around them. The paucity of their comprehension can best be demonstrated by examples of their responses. As an illustration of a social situation that posed no problem to the nonretarded eighth-grader , TSI picture #25 is illustrated as Figure 2.3. In interpreting the scene here presented, the average NMR eighth grader notes the cues: the turtle (or box of snakes), the facial expressions, the mother's gesture, the cat. He makes a series of rapid inferences: a boy, a woman with a cat--she is

---

[15] At least two of the "NMR" pupils had IQ's below 80 and six had IQ's between 80 and 90.

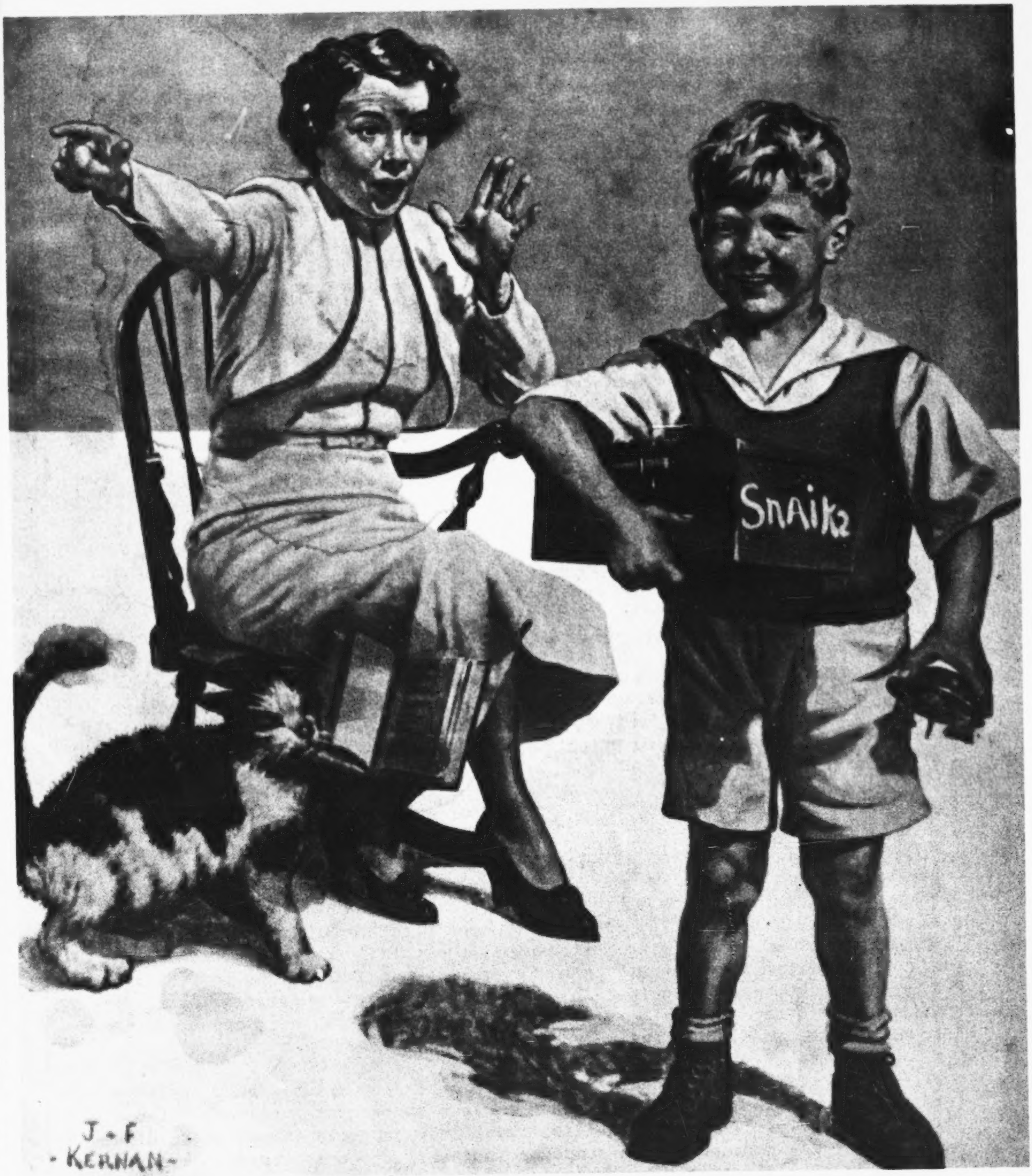

SnAIk2
J. F
-KERNAN-

Fig. 2.3 Picture #25 - Test of Social Inference

probably not a teacher because of the cat--she must be his mother. She looks upset; she is pointing and telling him something. The subject notes the reptiles and interprets the woman's gesture together with her expression, as telling the boy to get those things out this very minute. Instead of being displeased, the boy, however, looks happy. The nonretarded eighth-grader has no difficulty in inferring that the woman's emotion is the cause of the boy's pleasure. Typically his response consists of three or four inferences somewhat as follows: He brought in his pets to show his mother. She doesn't like them one bit/ she is upset by them. She is telling him to take them out this very minute. He thinks it very funny.

In contrast, responses to the same situation by low TSI performers among EMR subjects frequently suggest no awareness of the relationship between the boy and the woman, no awareness of the antecedent act of the boy's bringing in his pets, and, most surprisingly, no apparent comprehension of the woman's facial expression or gesture. Examples listed below are the interpretations of this picture by six pupils tested among the experimental population.

A lady is looking at the boy. The boys is looking at the cat.

There's a cat, a lady, a turtle. The boy is going outside to play.

He has a turtle. She dropped her book and tells him something. The cat is mad.

The boy has a turtle. Mother is pointing something out to him.

He brought in a turtle. He wants to play with it. His mother wants him to clean up his room. She dropped her book. The cat's watching.

He's going to get into trouble. Maybe the cat's not supposed to be in the house.

Don't know.

The responses quoted, although more typical of responses from the low performing institutional pupils, are actual responses by pupils enrolled as educable retarded in junior high school special education classes. Such examples appear to identify subjects in need of special rehabilitative effort. Their lack of awareness of relationships, their failure to comprehend the signal properties of the cues that are normally used as guides to behavior, demonstrate how poorly equipped they are to function without prosthetic assistance. In the absence of remedial training such subjects appear destined to an institutional or a sheltered workshop existence.

2(a). TSI Performance in Relation to Age: The relationship of TSI scores to age was examined using the same 24 classes and groups involved in the Table 2.4 analysis. The expectation that older subjects should score higher on the TSI was based on the assumption that ability to interpret the social situation is a function of learning; hence that the longer experiental history of an older subject favored improved TSI performance. As a control for gross differences between settings, examination of the TSI:age relationship was made separately for the NMR, the PS-EMR, and the Inst-EMR groups. The distribution of mean TSI scores together with the ranges of scores for each of these three groups according to successive age categories is presented in Table 2.6. As in Table 2.4, the data are compressed into summaries for pupils in each of the three groups, each section of columns in Table 2.6 representing the total number of pupils in that group.

TABLE 2.6

Distribution of TSI Scores Obtained by 88 NMR, 163 PS-EMR, and 108 Inst-EMR Pupils According to Age Categories

	TSI Score NMR			TSI Score PS-EMR			TSI Score Inst-EMR		
Age Categories	N	Mean	Range	N	Mean	Range	N	Mean	Range
228-251	-	-	-	1	70	-	14	36	18-78
204-227	-	-	-	16	64	18-98	38	43	5-82
180-203	41	81	50-108	62	58	18-98	36	42	6-84
156-179	47	84	53-116	70	51	9-104	15	31	5-60
132-155	-	-	-	14	50	8-72	4	25	16-31
below 132	-	-	-	-	-	-	1	24	-
Total	88	83	50-116	163	55	8-104	108	39	5-84

As may be noted from Table 2.6, the expected relationship of increasing TSI scores for increasingly older groups of pupils is in general maintained over the age ranges tested. The exception to the trend among the NMR may be an artifact related to the greater social perceptual deficit, together with older age, of one of the three NMR classes tested. The exception to the trend among the Inst-EMR is possibly explained by the movement from the institution of more socially competent older subjects back into community life. The TSI: age relationship, as indicated by Pearson product moment correlations, ranged for individual groups from -.36 to .38, with a median r of .19.

2(b). TSI Performance in Relation to IQ: The TSI:IQ relationship was similarly examined using the same 24 classes and groups of nonretarded and retarded pupils reported in Tables 2.4 and 2.6. The anticipated positive relationship between TSI score and IQ again related to the assumption of modifiability of social cue inference performance. In this instance, the assumption was that more generally intellectually handicapped persons would have had less social exposure and would have been slower in their learning to interpret social cues. Again, in partial control for setting, TSI:IQ relationship was examined separately for the PS-EMR and Inst-EMR samples. The distribution of TSI scores for these two groups is presented in Table 2.7 in terms of the mean and range of initial TSI scores for those pupils within each successive IQ category. As in the preceding Table 2.6, the data are compressed into single summaries for all pupil scores in these two groups, each section of columns in Table 2.7

representing the total number of pupils in that group.

TABLE 2.7

Distribution of TSI Scores Obtained by 41 NMR, 163 PS-EMR and 107 Inst-EMR Pupils According to IQ Categories

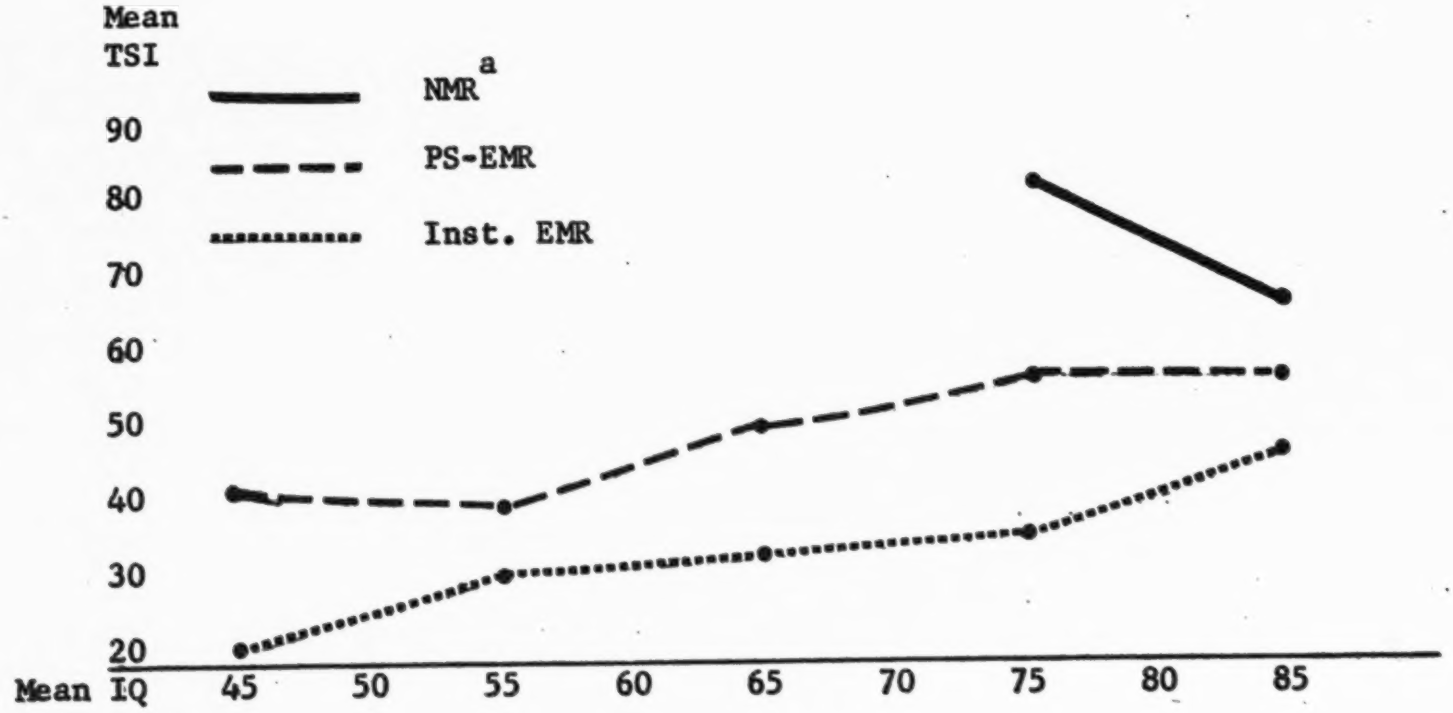
	TSI Score NMR			TSI Score PS-EMR			TSI Score Inst-EMR		
IQ Categories	N[a]	Mean	Range	N	Mean	Range	N	Mean	Range
80+	39	72	53-108	15	66	34-104	8	61	32-82
70-79	2[b]	78	68-89	68	60	8-98	24	44	14-84
60-69	-	-	-	54	54	18-90	35	42	8-78
50-59	-	-	-	23	39	9-77	31	30	5-58
40-49	-	-	-	3	43	39-50	9	22	10-38
Total	41	72	53-108	163	55	8-104	107	37	5-84

---

[a] IQ's not available for 47 additional NMR

[b] In spite of low IQ's, these pupils were attending regular 8th grade and are here termed NMR.

As can be noted from the Table 2.7, the expected relationship is well maintained in the EMR groups. The single exception from the trend of decreasing TSI means in relation to lower IQ category is based on a group of only three subjects having an IQ of below 50. The fact that for all IQ categories the TSI Means for the Inst-EMR pupils were always lower than the means for the PS-EMR pupils in the same IQ range is more exacting evidence of differences between institutional and noninstitutional retarded than the more gross data in Table 2.4. The Table 2.7 data can be seen more clearly in graphic form in Figure 2.4. Here the mean TSI score is shown in relation to IQ category separately for NMR, PS-EMR and Inst-EMR.



NMR N=	0	0	0	6	6 =	8
PS-EMR	3	23	52	70	15 =	163
Inst-EMR	9	31	37	22	8 =	107

[a] Of the 41 "NMR" pupils for whom IQ's were available, the IQ of two pupils was below 80; the IQ's of six pupils were between 80 and 90. The mean TSI score of the below 80 pupils is indicated at the IQ midpoint of 75. The mean TSI scores (65.3) of the between 80-89 pupils is indicated at the IQ midpoint of 85.

Fig. 2.4 Mean TSI Scores for IQ Intervals for 8 NMR, 163 PS-EMR, and 10 Inst-EMR Subjects

Only eight of the "NMR" pupils for whom IQ's were available had IQ's within the range depicted by Figure 2.4. Their TSI score inclusion provides tentative points of reference for the EMR performance. Although, as is illustrated in the figure, the PS-EMR and the Inst-EMR groups are clearly differentiated by mean TSI scores, it is worthy of note at the higher IQ end of the continuum that the 15 special education pupils whose average IQ was somewhat above the range that customarily delimits "retardation" had TSI scores lower in relation to IQ than would be expected from the relationship noted for the lower IQ pupils. One can only speculate whether their social comprehension and behavior may have been factors involved in their placement in the special education class. To extend the speculation into an examination of the NMR performance, perhaps the characteristics which "saved" the below 80 IQ subjects from special class placement were associated with their TSI performance that was close to the NMR average.

Of the seven Inst-EMR subjects whose IQ's were above 80 and whose average TSI score was higher in relation to IQ than would be expected from the relationship noted for the lower IQ pupils, six were girls whose appearance, conduct, and social competency were generally good. In certain instances their institutional placement was a means of removing them from unsatisfactory home environments.

The relationship between TSI scores and IQ's as indicated by Pearson product moment correlations, ranged for individual groups from -.06 to .78 with a median r of .60.

3. TSI Validity (Behavior Ratings): As noted earlier in this chapter, in terms of the TSI rationale it is assumed that a subject must acquire cultural categories for the classification of situations if his behavior is to be consistently situation-appropriate. He must also be able to identify the particular cues or attributes which may in effect be "read" as signalling or as standing for associated characteristics. Illustrative examples were given in Section B and in subsection F(1) of this chapter. The selection of TSI pictures in terms of their interpretability to an average eighth or ninth grader was assurance that social cue interpretation would be measured to some extent by the test. Legitimate questions, however, concern the extent of relationship between verbal responses given in an unhurried, simplified, picture-oriented, situation and the multi-modal responses appropriate to the changing complexity of a live situation. Although the testing of cue interpretation has its value as an indication of the extent to which a subject has or has not acquired this functional skill, the project's broader concern was with the subject's inerpersonal response; hence the need to determine the extent of relationship between TSI performance and more complex behavior.

For the purpose of assessing the more complex social behavior, definitions of social adequacy having apparent relevance to post-school situations as well as to pupils' contemporary milieux were developed for use in a series of rating scales (T-BR). One indication of adequacy, social range, was postulated as ability to move easily into a variety of social settings and to function easily within them. In closer focus upon interpersonal relations, two scales, social relationship and peer acceptance, were included, the one emphasizing transactional initiative and skill, the other focusing on the degree of peer acceptance. Then, as a kind of overview comparison of the ratee with others, a definition of social invisibility was formulated, indicating the extent to which the ratee's behavior was indistinguishable from that of nonretarded persons. An attentiveness scale was included as being related to the project's interest in social perception. Four additional criteria were added to avoid biasing teachers in favor of particular forms of adequacy. The added scales required ratings of academic skill, calmness, industry and appearance as defined by Parnicky and Kahn (1963). A similar, although modified, set of scales (O-BR) was prepared for use by a group of nonteacher raters. These scales, both T-BR and O-BR, are described and their interrelationships discussed in Chapter III.

Examination of the relationship between TSI scores and the teachers' ratings (T-BR) was possible for each of the four experimental trials of the training curricula. These trials involved 9 classes of PS-EMR pupils and 6 classes of Inst-EMR pupils. In all, 115 PS-EMR pupils, 69 boys and 46 girls, and 70 Inst-EMR pupils, 32 boys and 38 girls, were included in these trials. These 15 classes are among the 21 classes of retarded subjects reported in Tables 2.4 and 2.6.[16] Behavior ratings and pupil TSI performance scores were obtained for each pupil immediately preceding and immediately following the periods in which the social perceptual lessons were used experimentally, as described in Chapter V. For the ensuing discussion of TSI-TBR relationship, only the pre-training data are used.

Pearson product moment correlations were computed between TSI scores and teacher ratings on the nine T-BR scales, combining all of the girl pupils into one sample of 84 pupils. The boys were similarly combined into one sample of 101. The coefficients are presented in Table 2.8.

TABLE 2.8

Correlation Coefficients of TSI Scores with Initial Teacher Ratings on Nine Scales for 84 Girls and 101 Boys

		Social Scales				
Sex	N	Social Range	Social Relat.	Peer Acceptance	Social Invis.	Attentiveness
Girls	84	.59**	.47**	.52**	.15	.26*
Boys	101	.36**	.29*	.29*	.12	.30*
Combined	185	.46**	.36**	.39**	.12	.26**

		Additional Scales			
	N	Industry	Acad. Skill	Calmness	Appearance
Girls	84	.17	.15	.38**	.12
Boys	101	.17	.19	.23**	.05
Combined	185	.14	.14	.30**	.08

* < .05 l.c.
** < .001 l.c.

In the upper portion of the table the correlations shown are those between TSI and the hypothesized relevant "social" scales of social range, social relationship, peer acceptance, social invisibility and attentiveness. TSI scores are shown to be moderately correlated with T-BR ratings on four of the five scales, all correlations attaining the .05 level. However, the low positive correlation between TSI scores and ratings on social invisibility does not attain this level of significance.

---

[16] The ages, IQ's and TSI scores for these pupils are listed separately in Tables 5.1 and 5.2, 5.13 and 5.14 in Chapter V.

In the lower portion of the table the correlations shown are those between TSI and the scales expected to be less closely associated. As hypothesized, the low positive correlations between TSI scores and T-BR ratings on industry, academic skill, and appearance do not attain the .05 level. The moderate correlation between TSI scores and ratings on calmness, however, is of approximately the same magnitude as the TSI:social scale correlations. TSI scores are shown by this data to be most closely related to the kinds of social adequacy rated as social range, peer acceptance, social relationship, attentiveness and calmness.

Examination of TSI scores in relation to ratings by the outside (nonteacher) raters (O-BR) was possible for the two Trial 3 classes, involving 34 institutional pupils, 14 boys and 20 girls, for whom ages and IQ's are included in Tables 2.5 and 2.6. Pearson product moment correlations were computed relating the pupil's TSI score to the average rating he received on the O-BR scales from six raters. These coefficients are presented in Table 2.9.

TABLE 2.9

Correlation Coefficients of Initial TSI Scores with Mean Initial Ratings by Six Outside Classroom Raters on Nine Scales for 20 Inst-EMR Girls and 14 Inst-EMR Boys

		Social Scales			
Sex	N	Social Range	Social Accept.	Social Invis.	Attentiveness
Girls	20	.75**	.63*	.31	.52*
Boys	14	.48	.51	.45	.38
Combined	34	.65**	.56**	.32	.42*

		Additional Scales				
Sex	N	Industry	Acad. Skill	Calmness	Appearance	Honesty
Girls	20	.25	.45*	.51*	.07	.18
Boys	14	.31	.39	.39	.40	.37
Combined	34	.22	.37*	.42*	.15	.24

---

* < .05 l.c.

** < .001 l.c.

In the upper portion of Table 2.9 are listed correlations between TSI scores and ratings on the "social scales." The definitions for two of the T-BR scales, social relationship and peer acceptance, were reformulated so as to combine the two highly intercorrelated T-BR scales into a single O-BR scale, social acceptance. In this sample of a single-institutional group, TSI scores were found to be most highly correlated with ratings on social range, social acceptance, attentiveness, calmness, and academic skill. The relationship was higher for girls than for boys. TSI scores were found slightly less associated with social invisibility, industry, appearance and honesty. The differentiation between TSI:"social scale" correlations and the TSI:"additional scales" correlation that can be noted in the previous T-BR data in Table 2.8 is less sharp here. The explanation appears to reside in the high intercorrelations between O-BR scale mean ratings.[17]

To carry the investigation of the meaningfulness of TSI scores a step beyond the preceding discussion of their positive relation with IQ alone, and with behavioral ratings, alone, a comparison of the IQ:T-BR correlations with TSI:T-BR correlations is shown graphically in Figure 2.5 for the 84 EMR girls and 101 EMR boys for whom data was presented in Tables 2.8 and 2.9.

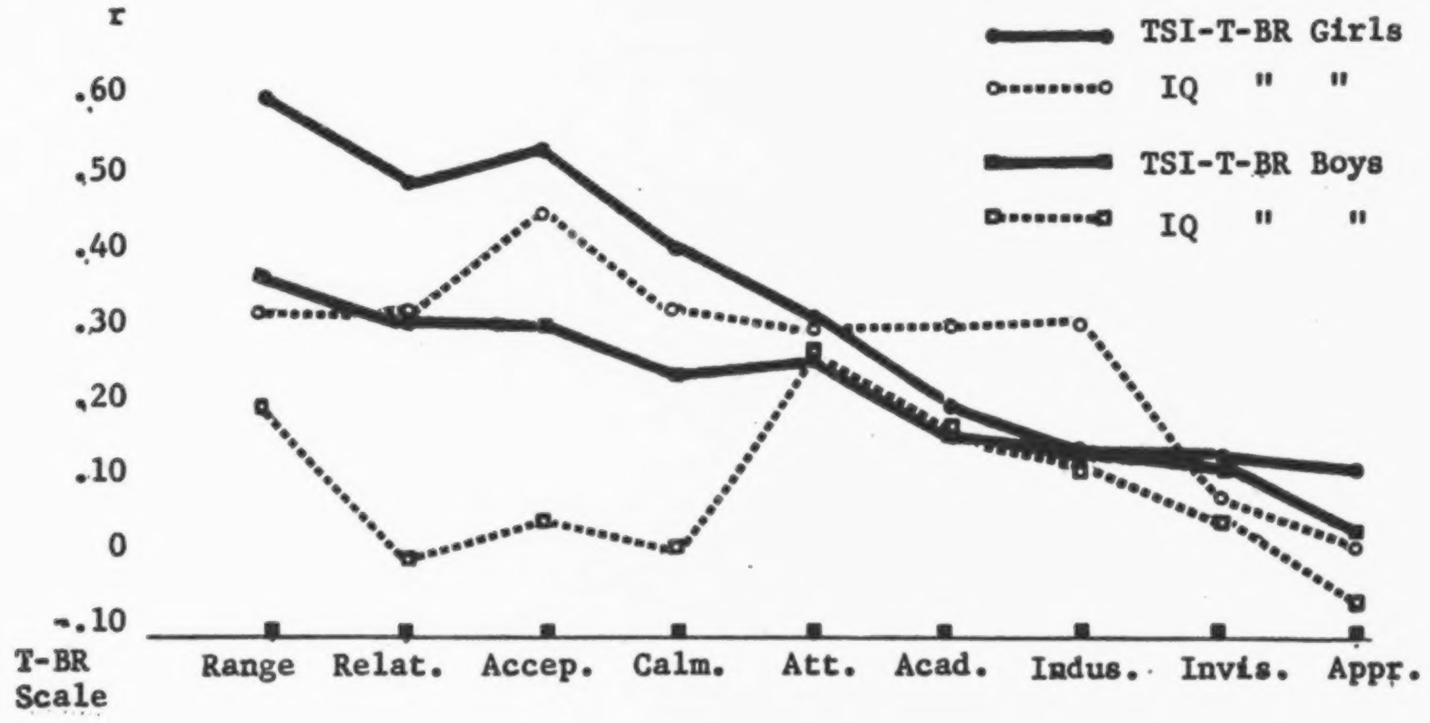


Fig. 2.5 Pearson Product Moment Correlations for 84 EMR Females and 101 EMR Males between TSI scores and T-BR Ratings on Nine Scales and between IQ's and T-BR Ratings.

---

[17] A discussion of the O-BR scale interrelationships is presented in Chapter III (p. 3.18)

In this figure the vertical axis represents Pearson product moment correlations while points on the abscissa represent the nine T-BR scales. To make the differentiation between TSI and IQ clearer, the correlations involving TSI scores are joined by continuous lines; the correlations involving IQ are joined by a dotted line. The figure reveals that for both sexes IQ's tend to parallel TSI scores in their relation to T-BR ratings of behavior attributes but that the relationships diverge in relation to the scales social range, social relationships and peer acceptance. For both sexes TSI scores were more closely associated with the behavioral characteristics rated on those scales than were IQ's. For girls, ratings on the social scales and on calmness were more highly correlated with both TSI scores and IQ's than was true for boys. For girls, ratings on academic skill and industry were somewhat more closely associated with IQ than with TSI score. For boys this difference was not obtained. For both sexes the TSI:attentiveness correlations and the IQ: attentiveness correlations were practically identical. In addition to illustrating TSI score validity in its relation to ratings of social adequacy, the figure demonstrates that neither TSI score nor IQ's are related to ratings of social invisibility, a finding discussed in section C of Chapter III.[18]

4. TSI Validity (TSI in Relation to Other Criteria): Data available from a study by Clark (1967) is also relevant to discussion of the meaningfulness of TSI scores. In an independent partial replication of the project procedures, Clark's sample consisted of 56 EMR pupils enrolled in six special education classes at junior high school level in Nashville Public schools. Thirty-five were male; 21 were female. Their ages ranged from 13 to 16 with a mean age of approximately 15. IQ's ranged from 53 to 80 with a mean of 68. In connection with his trial of the experimental curriculum II developed by the present authors, Clark administered the Vineland Social Maturity Scale, the Fudell Test of Occupational Readiness, and theBown Self Report Inventory. He also used rating scales similar to the O-BR format of the present project, pupils receiving pre- and posttraining ratings by their teacher and by other official observers.

The Vineland Social Maturity Scale (VSMS) is a behavior check list to be filled out by a rater well enough acquainted with the ratee to be capable of indicating specific competencies such as "talks in short sentences"; "relates experiences"; "goes about neighborhood unattended"; "does small remunerative work"; etc. Its subscales include items that are components of Self-Help, Self-Direction, Locomotion, Occupation, Communication, and Social Relations. Although ratings are most often made by a parent or a teacher, subjects who are capable of self-appraisal may rate themselves on the instrument.

---

[18] The social invisibility scale's apparent independence of IQ and its moderate to high association with ratings on attentiveness, academic skill, industry, social relationship, calmness and appearance suggests its possible utility in predicting a successful transition from school or institution into the adult world.

As used on the Clark study, pupils made self-ratings on the VSMS.[19]

The Fudell Test of Occupational Readiness (FTOR) was described by Clark as an achievement measure to accompany Fudell's vocational prepraredness lessons (Fudell, 1963) in programs for EMR pupils at secondary school level. The test was designed for group administration when pupils are able to read the questions; comprehend them, and respond by underlining either Yes or No. Fudell devised and used the test to assess the extent to which a pupil had acquired basic information and attitudes deemed essential for entrance into an occupation. No reliability coefficients are reported.

The Bown Self Report Inventory (BSRI) adapted by Stephens for retarded persons was designed as a measure of a subject's self-perception and of his perceptions and feelings about children, authority, work, parents, and the future. When administered to the retarded the questions are read aloud to the subject and his reply is indicated by his marking either Yes or No responses. A test-retest reliability coefficient of .88 was reported for the testing of 25 subjects (Peck, Stephens, & Fooshee, 1964).

The four teacher Behavior Rating Scales (T-BRS) and the four Outisde Rater Scales (O-BRS) as used by Clark were five point graphic scales similar in format to the outside rater form (O-BR) developed by this project and described in Chapter III. The scale criteria, Social Range, Social Acceptance, Social Invisibility, and Attentiveness were the same as those described in Subsection F(3) above. In reporting his data, Clark averaged the ratings made on all four scales for a composite rating of pupil adequacy.

Correlations between pupils' scores on each of these instruments with TSI scores, and the relationship of each instrument with the other measures are presented as a matrix in Table 2.10

TABLE 2.10

Correlations between Scores on TSI, FTOR, BSRI, and Composite Behavior Ratings on Four T-BRS and O-BRS Scales for 56 EMR Pupils

	TSI	VSMS	FTOR	T-BRS	O-BRS	BSRI
TSI	-	.51*	.51*	.27*	.41*	.21
VSMS		-	.23	.10	.32*	.21
FTOR			-	.35*	.41*	.14
T-BRS				-	.25	.00
O-BRS					-	.09

* < .05 level

---

[19] Doll (1953, p.431) reported a reliability coefficient of .94 between two administrations of the VSMS to 48 retarded subjects when ratings were made by adults. Use of the VSMS for self-ratings by retarded subjects (N=67) was reported as being nearly as reliable (r=.90) but as yielding scores somewhat higher.

From the coefficients in Table 2.10 it is evident that TSI scores were moderately correlated with pupil self ratings on the VSMS, with pupil scores on the FTOR, and with the composite social adequacy ratings by teachers and other raters. Pupil scores on the Test of Social Inference were related to their social adequacy ratings to approximately the same extent as were their scores on the Vineland Social Maturity Scale and their scores on the Fudell Test of Occupational Readiness. The Bown Self Report Inventory, although positively correlated with TSI and VSMS, was relatively independent of other measures.

Summary

Hypothesizing a relationship between a retardate's social behavioral inadequacies and his failure to interpret the social scene, the goal of the project was to increase his social adequacy by teaching him to make inferences from social cues. The Test of Social Inference (TSI), developed for use as a pre- and posttraining measure of the adolescent's comprehension of visual cues, is described in this chapter.

In addition to its description of TSI rationale and administration, the chapter presents test-retest and interscorer reliability data. Reliability coefficients in the form of Pearson product moment correlations ranging from .84 to .97, with a median r of .91, were obtained from two administrations of the TSI to groups of EMR pupils. The validity of the TSI is discussed in terms of the relationship between TSI scores and subject characteristics such as age, IQ, place of residence, ratings given on scales of social adequacy, and scores earned by pupils on other instruments. Pearson product moment r's ranging from .26 to .71 were obtained for the relationship between TSI scores and teacher ratings on the scales social range, social relationship, peer acceptance, attentiveness and calmness. TSI scores were found moderately correlated with scores by 56 pupils on the Vineland Social Maturity Scale, $r=.51$ ($p= <.05$) and with scores on the Fudell Test of Occupational Readiness, $r=.51$ ($p= <.05$)

Mean TSI scores of NMR, PS-EMR and Inst-EMR subjects differed from one another at less than the .05 level. The average TSI scores of PS-EMR pupils differed from Inst-EMR pupils when IQ's were held constant.

III. THE RATING OF SOCIAL BEHAVIOR

In a generic sense this project was initiated in response to the problem of the asocial and maladaptive behavior of the retarded. Congruent with this concern, the project schedule provided for the development of an experimental training curriculum for the improvement of the retardate's social behaviors and for the trial of these materials in various educative settings. In part, the development of the Test of Social Inference (TSI) described in Chapter II was compelled by a need to measure the effectiveness of this experimental curriculum. But the TSI as a contrived "laboratory" test of social cue interpretations affords, at best, only partial assessment of curricula designed for behavioral change. A fuller report of what the posttraining subject (and his control counterpart) does is needed to determine how, if at all, the behaviors to be modified by the training have been changed.

This chapter describes project activities initiated by this concern for more adequate measurement of behavior change. The activities center around the development of workable techniques for observation and rating of behavior. The first section of the chapter describes trials that were made of direct observation and recording techniques in naturalistic situations. In a following section are presented the two rating scale formats developed in this project, the Teacher Behavior Rating Scale (T-BR) used in all of the experimental curriculum trials as a pre- and posttreatment assessment of effect, and the Outside Rater Behavior Rating Scale (O-BR) used in only one of the reported curriculum trials.

In an additional section the concept of social behavioral adequacy is discussed through consideration of the intercorrelations of ratings that were obtained in the course of the project by use of the T-BR and O-BR scales.

In a last section concerned with the meaningfulness of the behavior data, there is discussion of the differing discriminations required by the rating task and a discussion of the teacher's and outside rater's behavior as a rater.

A. Direct Observation of Behavior

Although the use of behavioral rating scales had been originally proposed as the method of assessing the hoped-for social behavioral gains, it was recognized that rating scales provided less precise information than would some method of actual direct observation of pupils.[20] In part because patterns of socially appropriate behavior differ in accordance with setting, there is risk in generalizing from observations in only one setting. Because of project concern for the postacademic adjustment of the retarded pupil, there was particular interest in his behavior in non-academic situations. Ideally to be desired would be a method of direct observation of behavior in a number of settings, not excluding the academic. Efforts toward this end extended over several months, during which time recommended observational techniques were tried out. In the most successful of these, two observers recorded 20-second samples of behavior for each member of a pupil group, the observers following a pre-arranged sequence for their observations. Simultaneity of observation by the recorders was synchronized by a device that clicked softly at 20-second intervals. The recording of each behavioral sample observation was prefaced by the recording of the contextual activity or type of setting in which the behavior was taking place. Two trials of the method yielded satisfactory inter-observer reliability coefficients. Interjudge agreement as to the social appropriateness of the recorded behaviors was likewise attained. There remained, however, both immediate and more general problems that counter-indicated the use of direct observational methods. The immediate impracticality was the inexorable project deadline. For the experimental materials and methods to be given classroom trials, all procedures, including the observational method, had to be operational a sufficient time before the end of a school year to insure the 8 to 10 weeks necessary for the pre- and posttesting and for the experimental training trials. The longer range impracticality was budgetary. The observational trials demonstrated that to make such detailed observations is expensive. Much personnel time goes into practice. Then the recordings must be transcribed. There then remains the process of rating, scoring, or of evaluating the transcribed behaviors as to appropriateness within the context of the situation. For these reasons of time and expense, the project focus returned to the use of behavioral rating scales and to the classroom teacher as a most convenient and, in many ways, well qualified observer.

---

[20] The distinctions made here between "direct" behavior observations and use of teacher or outside rating scales is more correctly one between ratings based on immediately recorded observations and ratings based principally on remembered, nonrecorded data. All ratings used in projects are based on personal rater observation. The teacher and outside ratings are less direct only in the sense of not being made with more objective reference to descriptions of the S's behavior during the observational period.

In giving up efforts at direct observations and in reverting to behavior ratings to be made by classroom teachers, the desire was not abandoned to assess non-classroom behavioral adequacy. Although the teacher has less contact with his pupils during outside classroom activities, limiting his cognizance of more varied pupil performances involved in social behavioral adequacy, studies have indicated the value of teacher ratings (Ullman, 1957; Voelker, 1962). The problem, of course, is one of available observer-raters of the non-classroom setting--observers personally knowledgeable about all of a sample of pupils. In the current situation, as the best available source of information, the teacher was used as a behavior rater in each of the treatment trials. In addition to the teacher as a rater, at one of the state institutions cooperating with the testing of the trial curriculum, six persons were found well qualified as raters because of their acquaintance with all of the pupils to be rated. These were the chaplain, nurse, librarian, school secretary, director of cottage life and recreation director. As will be described, the rating score format and procedures used by the outside raters represented some improvement upon the format and procedures employed by the teacher raters.

B. Description of Teacher and Outside Rater Behavior Rating Scales

(1) The Teacher Behavior Rating Scales developed in this project use a five-point graphic rating scale format similar to the Vocational Performance and Personality Adjustment Scales employed in a study reported by Parnicky and Kahn (1963). Several modifications, however, were made in the scale format. Under the original format, a subject was apparently rated on each of the scales before the rater proceeded to rate the second subject and so on. For use in the current project, the rating forms were changed so that each subject in a group would be rated with respect to each second criterion. The rating page was divided into five columns with a listing of pupil names at the left margin. The columns corresponded to the five alternative ratings of "very poor," "poor," "fair," "good," and "very good." Each alternative rating was defined in terms of a brief description of behavior characteristics of pupils belonging in that column. The rating required a matching of pupil and behavior description to be indicated by a check mark in the appropriate column.

Because of the project interest in social performance over a broad range of settings, definitions of behavioral adequacy were developed which appeared relevant both to present pupil life and to postschool society, each definition, in effect, constituting a criterial scale. The social range scale focused on the subject's ease of movement in and out of a wide range of settings. The social relationship and peer acceptance[21] scales both focused on social interaction, the former placing the greater emphasis on the ratee's approach and transactional initiative and skill, while the latter emphasized peer response. The social invisibility[22] scale was intended to assess the extent to which the retardate's behavior was indistinguishable from that of nonretarded persons. Because of project interest in social perception, a scale to assess a subject's attentiveness was added. This attentiveness scale focused on the subject's attention span and general attentiveness. These five scales are included as Appendix .

In addition to the four "social adequacy" scales and the attentiveness scale, four additional criteria from the Vocational Performance Scales (Parnicky & Kahn, 1963) were added to the T-BR Scales as a way of de-emphasizing, during the experimental trials, project interest in particular aspects of behavior. To have provided only the four social scales mentioned above might have biased the teachers in their judgment of the curricular lessons by suggesting the project emphasis. The added scales required teacher ratings of the pupil's appearance, industry, calmness, and academic skill. Definitions of these scales are included in Appendix B . The added scales were selected as representative of additional behavioral characteristics of relevance both to the academic and non-academic milieux. These characteristics, however, were considered tangential to the present project focus on social cue interpretation.

---

[21] As used in the study, the scale was named social acceptance. It is here renamed peer acceptance, as better denoting its criterion.

[22] A fuller discussion of the relevance of social invisibility to the retarded is provided in Leland & Smith (1965, Ch. 10, pp 31-33)

(2) A second behavior rating form, O-BR, was used in connection with one of the experimental trials of the training curriculum. In an institutional setting (described as Trial 3 in Chapter V) several members of the institutional staff were found to be sufficiently acquainted with the experimental and comparison group subjects to serve as raters of the pupils' behavior in settings outside of the classroom. The behavioral definitions provided these "outside" raters were, with two exceptions, closely similar to the behavioral definitions provided for the teacher raters in the T-BR scales. As one modification, the T-BR social relationship and the peer acceptance scales were merged into a single scale of social acceptance. The merging was dictated by the high correlations between these two scales. As a second modification, in replacement of the scale thus eliminated, an honesty scale was added to the O-BR.

An improvement upon the T-BR format was introduced in response to rater use of limited portions of the scale. As is discussed in greater detail in the following section, some teachers had been noted to restrict their ratings to the low portion of the scale; some others restricted their ratings to one or two middle categories. In redrafting the scale format an explicit frame of reference was introduced in an attempt to have raters make better use of the full scale. The idea of providing the rater an explicit frame of reference originated with Scott (1941) and was more recently reintroduced by Gardner and Thompson (1956) in developing their Syracuse Scales of Social Relations. In following the Gardner and Thompson procedures, actually the rater provides his own frame of reference aided by the instructions and a supporting format on the rating form. As a part of this format, at the top of the rating form, five boxes are provided for the rater's use in filling in the names of the persons who represent to him, respectively, the most outstanding individual he can recall with respect to the listed criterion, the lowest, the merely average, the halfway between highest and average, and the halfway between lowest and average. The rater is required to develop a referent set of names, in this manner, for each scale.

The instructions used in the present study were similar to those developed by de Jung and Haring (1962) in their modification of the Syracuse Scales. These required the rater to consider all retarded persons he could recall of approximately the same age as the pupils to be rated. From this recalled population he was instructed to select five referent persons with respect to the given criterion: a most outstanding or highest person, a "least" person, a "middle" person, and a "between most and middle" and a "between least and middle" person. The names of these five individuals were to be written into the separate boxes, labeled "least" to "most", at the top of the rating page. To facilitate the rater's recall of appropriate reference persons, raters were instructed to engage in a pre-rating oral period of discussion and recall of retarded persons they had known of approximately the age of the pupils to be rated.[23] Immediately following the recall period the raters filled in their frame of reference boxes independently and proceeded with their ratings of the pupils whose names were listed on the rating form.

The O-BR rater instructions and reproduction of the O-BR format, together with each of the scale criteria, are included as Appendix C.

C. Examination of Teacher Behavior Rating Scales (T-BR)

During the course of project activities repeated teacher ratings were made of 15 classes of retarded pupils. These 185 pupils were involved in the four trials[24] of the experimental social cue training curriculum. These consisted of five classes of EMR adolescents, aged 12 to 17½, enrolled in special education classes in the Kansas City area public schools; six classes of pre-vocational level institutionalized retarded youngsters, aged 12 through 20, from the Parsons State Hospital and Training Center and from the Hissom Center, Sand Springs, Oklahoma; and four from further classes of public school special educational class EMR adolescents enrolled in Springfield, Missouri and Kansas City public schools, aged 12 ½ to 17. These fifteen classes of pupils are more fully described as classes A through O in Chapter V of this report. In all classes pre-experimental treatment and posttreatment ratings were obtained on all pupils. In the first five classes, (A-E) the teacher ratings were made weekly for 13 weeks, with the pre- and postexperimental rating computed as averages of the first three weeks and of the last three weeks of teacher ratings, respectively. In the first institutional setting (classes F-I) the rating period paralleled the experimental curriculum period and the pre- and postexperimental ratings were computed as averages over the initial two weeks of ratings and over the final two weeks of ratings, respectively. In the second institutional setting (classes L-O) only one weekly rating was made the week prior to the experimental curriculum program and the week following its termination, these single ratings serving as the pre- and postratings.

The T-BR ratings were obtained for two purposes: one use was to obtain a measure of pupil social adequacy before and after the experimental training, permitting an assessment of treatment-associated behavioral change, and to permit examination of the postulated relationships between ratings assigned pupils on the social scales, social range, social relationship, peer acceptance, and social invisibility, and on the attentivenss scale, with their scores on the Test of Social Inference. The T-BR data relevant to these two foci are discussed in Chapters II and V of this report.

---

[23] In terms of the rationale, the preliminary period of recall and selection of a representative frame of reference should lead to more objective ratings. A further advantage argued for this procedure is interrater comparability of ratings. (de Jung, 1964)

[24] These trials are separately described in Chapter V.

A more general consideration is that of the meaningfulness of these ratings. How does a rater, in this instance, a teacher, use the scales? (a) How well does he discriminate between the behavioral characteristics he has been asked to rate? (b) How consistent are his ratings? (c) How well can he use the rating scales to assess change in behaviors? Within the present project content, these are all questions concerning the teacher as a rater--can he do what is required of him? These are questions both of relevance to the general problem of behavioral assessment and, more immediately, to interpretation of the experimental results of the project. Although these questions cannot be fully resolved in terms of data from the current project, the following analyses of the 15 class rating data afford some preliminary answers, particularly as related to the project concern with assessment of treatment-associated behavioral change.

(a) A first examination of teacher discrimination in the use of the several rating scales was made in terms of an analysis of the variability in ratings received by a pupil on the several rating scales. This analysis was made of the pre-treatment ratings. To afford more direct relevance to the actual rating data used in describing pupils' social behavior, average ratings[25] rather than single ratings were used. For each pupil, the range of ratings received by him on the nine scales was recorded as a ratio of maximum possible scale range; namely, five scale units. If a pupil received ratings of only one through three on his nine scales, his range was .6 of maximum. To provide a summary description of the teacher's differential ratings, average ratios were computed for each teacher based on all those pupils he rated. The teacher averages together with the minimum and maximum range ratios computed for his pupils are reported in Table 3.1

TABLE 3.1

Average Teacher Use of Scale Range in Rating Individual Pupils on Nine Behavioral Rating Scales

	Range of Scale Intervals Received by a Pupil on 9 Scales			
Teacher	Class Size	Aver. Range Ratio	Max.	Min.
A[a]	11	.65	5	2
B[a]	8	.50	3	2
C[a]	15	.59	4	2
D[a]	8	.53	3	2
E[a]	12	.68	5	2
F[b]	9	.44	3	2
G[b]	9	.49	3	2
H[b]	9	.51	4	2
I[b]	9	.47	3	2
J[c]	17	.69	5	2
K[c]	17	.81	5	3
L[c]	12	.55	3	2
M[c]	24	.56	4	2
N[c]	12	.63	4	2
O[c]	11	.67	5	3

a Pre-ratings were the average of 3 pre-treatment ratings
b " " " " " " 2 " " "
c " " " a single-pre-treatment rating

---

[25] For those nine classes having more than one set of pre-treatment ratings, the two or three sets of pre-ratings were combined into a single, average pre-treatment rating.

As can be noted in Table 3.1, the average range ratios for the 15 teachers varied from a low .44 (Class F) to a high .81 (Class K); in effect, from less than half to four-fifths of maximum discriminability. The median average for the 15 teachers was .56 of the scale range. Of the 15 teachers, five varied their ratings giving at least one of their pupils the full five point scale range; that is, rating a pupil "lowest" on one (or more) of the nine scales and "highest" on one (or more) of the remaining scales. Another four teachers used as many as four of the five scale portions in rating at least one of their pupils. Apparently the pupil is a factor in determining interscale discriminations on the part of the teacher. The summary picture seems to be one of satisfactorily broad use of the rating continuum for a given pupil rated on several scales, with some teachers being more discriminative and some pupils being more discriminatively rated on the different scale criteria. These differences with respect to teachers are not surprising especially with minimal rater training. The difference with respect to pupils is encouraging as support for meaningful teacher rating task behavior, in that the spread of ratings varies with the pupil rated and not as a forced artifact of the rating format and instruction.

A further examination of teacher rater interscale discrimination was made in terms of Pearson product moment correlations computed between ratings made on the nine different scales. As in the previous analysis, only pre-treatment ratings were examined. In computing these coefficients the fifteen classes were collapsed into three samples: five public school special educational classes (A-E) comprising those participating in the 1965 studies, six pre-vocational institutional classes (F-K) comprising those participating in the 1966 studies, and four public school special education classes (L-O) comprising those participating in the 1967 studies. A further subsampling was made within each of the three groups in terms of sex of ratee. This latter separation seemed warranted because of apparent teacher differences in ratings given boys and girls. The resulting six samples ranged in size from 22 to 38. In computing the six matrices of intercorrelations of ratings, three additional pupil variables were also included, pupil age, IQ, and TSI score. The coefficients are also presented, separately, by sex, for the combined samples. Table 3.2 presents these 864 correlations. To facilitate discussion of these intercorrelations, they are presented in nine groupings, one for each of the nine rating scales.

Table 3.2 about here

As may be noted in Table 3.2, the most consistently high correlations across samples are those involving scale 5, attentiveness, with scale 6, industry, r's of .76, .74, .70, .71, .80, and .88 (median .78 for the six samples respectively and with scale 7, academic skill, r's of .67, .46, .81, .74, .68, and .95 (median .71) respectively. The intercorrelations between these latter two scales, 6 and 7, are also consistently moderate to high; r's of .59, .66, .74, .68, .74, and .84 (median of .71). This intercorrelational cluster of coefficients involving attentiveness, industry and academic

TABLE 3.2

Interscale Correlations Based on Pre-Training Teacher Ratings of 101 Male and 84 Female EMR Pupils on Nine T-BR Scales

Scale No.	N	1	2	3	4	5	6	7	8	9	CA	IQ	TSI
1 Soc.Range													
Males													
1965	32	-	.73	.65	.35	.53	.35	.07	.39	-.10	-.20	.32	.43
1966	32	-	.64	.70	.51	.60	.61	.41	.74	.09	.08	.38	.37
1967	37	-	.68	.77	.08	.32	.26	.38	.32	.36	.05	.11	.30
Total	101	-	.71	.72	.34	.48	.38	.34	.57	.18	-.06	.18	.36
Females													
1965	22	-	.80	.74	.62	.74	.63	.54	.59	.65	-.04	.73	.70
1966	38	-	.53	.54	.11	.36	.25	.48	.38	-.06	.32	.52	.71
1967	24	-	.44	.46	.45	.35	.33	.41	.28	.33	.45	-.23	.29
Total	84	-	.60	.57	.38	.46	.33	.46	.46	.24	.30	.32	.58
2.Soc. Relat.													
Males													
1965	32	.73	-	.65	.54	.46	.54	.06	.41	-.07	-.41	.06	.38
1966	32	.64	-	.70	.81	.64	.83	.63	.87	.07	.28	.10	.33
1967	37	.68	-	.77	.32	.48	.46	.60	.64	.50	.14	.02	.25
Total	101	.71	-	.72	.58	.54	.58	.50	.72	.26	.06	-.01	.29
Females													
1965	22	.80	-	.84	.80	.78	.64	.51	.70	.48	.06	.54	.58
1966	38	.53	-	.68	.34	.56	.52	.42	.50	.17	.08	.39	.53
1967	24	.44	-	.86	.52	.29	.27	.31	.52	.45	.50	.14	.25
Total	84	.60	-	.76	.52	.52	.43	.53	.58	.34	.17	.30	.47
3.Peer Accept.													
Males													
1965	32	.55	.65	-	.46	.52	.16	.26	.16	.15	-.20	-.11	.23
1966	32	.50	.70	-	.48	.49	.56	.49	.56	.07	.18	.21	.23
1967	37	.51	.77	-	.44	.48	.64	.54	.64	.60	.26	.06	.39
Total	101	.53	.72	-	.47	.50	.50	.46	.50	.31	.06	.05	.29
Females													
1965	22	.74	.84	-	.80	.68	.74	.58	.74	.36	-.14	.66	.50
1966	38	.54	.68	-	.15	.34	.45	.46	.45	.12	.20	.53	.55
1967	24	.46	.86	-	.55	.35	.55	.30	.55	.60	.42	.11	.44
Total	84	.57	.76	-	.38	.44	.56	.40	.56	.30	.14	.42	.52

TABLE 3.2 - cont.

Scale No.	N	1	2	3	4	5	6	7	8	9	CA	IQ	TSI
4. Invisibility													
Males													
1965	32	.35	.54	.46	-	.59	.57	.31	.75	.33	-.18	.13	.15
1966	32	.51	.81	.48	-	.50	.58	.53	.74	.29	.24	-.04	.07
1967	37	.08	.33	.44	-	.72	.62	.67	.67	.60	.07	.07	.12
Total	101	.34	.58	.47	-	.62	.58	.53	.70	.43	.06	.02	.12
Females													
1965	22	.62	.80	.80	-	.84	.70	.64	.91	.42	-.04	.34	.39
1966	38	.11	.34	.15	-	.48	.36	.35	-.02	.58	-.17	-.05	.09
1967	24	.45	.52	.55	-	.81	.72	.85	.61	.69	.36	.27	.27
Total	84	.38	.52	.44	-	.69	.58	.58	.42	.59	.14	.06	.15
5.Attentiveness													
Males													
1965	32	.53	.46	.52	.59	-	.76	.67	.39	.24	-.06	.21	.33
1966	32	.60	.64	.49	.50	-	.74	.46	.71	.00	.26	.45	.42
1967	37	.32	.48	.48	.72	-	.70	.81	.50	.53	.26	.32	.29
Total	101	.48	.54	.50	.62	-	.72	.70	.54	.34	.14	.23	.30
Females													
1965	22	.74	.78	.68	.84	-	.91	.74	.78	.62	.01	.36	.41
1966	38	.36	.56	.34	.48	-	.80	.68	.44	.11	.22	.35	.32
1967	24	.35	.29	.35	.81	-	.88	.95	.37	.63	.21	.19	.22
Total	84	.46	.52	.44	.69	-	.84	.80	.50	.44	.26	.25	.26
6. Industry													
Males													
1965	32	.35	.54	.56	.57	.76	-	.59	.26	.15	-.16	.14	.17
1966	32	.61	.83	.63	.58	.74	-	.66	.78	.06	.10	.23	.35
1967	37	.26	.46	.46	.62	.70	-	.74	.35	.56	-.12	.14	.03
Total	101	.38	.58	.54	.58	.72	-	.66	.44	.30	-.04	.16	.16
Females													
1965	22	.63	.64	.56	.78	.91	-	.68	.73	.57	.02	.24	.20
1966	38	.25	.52	.37	.36	.80	-	.74	.55	-.02	.01	.42	.26
1967	24	.33	.27	.31	.72	.88	-	.84	.50	.50	.23	.27	.21
Total	84	.33	.43	.34	.58	.84	-	.76	.52	.33	.12	.31	.16

TABLE 3.2 - cont.

Scale No.	N	1	2	3	4	5	6	7	8	9	CA	IQ	TSI
7.Academic													
Males													
1965	32	.07	.06	.26	.31	.67	.59	-	.10	.25	-.04	.27	.14
1966	32	.41	.63	.49	.53	.46	.66	-	.58	.45	-.12	.26	.19
1967	37	.38	.60	.54	.66	.81	.74	-	.51	.68	.26	.30	.30
Total	101	.34	.50	.46	.53	.70	.66	-	.44	.52	.04	.17	.18
Females													
1965	22	.54	.51	.58	.64	.74	.68	-	.61	.42	-.15	.16	.20
1966	38	.48	.42	.46	.35	.68	.74	-	.35	.00	.23	.45	.28
1967	24	.41	.31	.30	.85	.95	.84	-	.37	.55	.21	.19	.08
Total	84	.46	.53	.40	.58	.80	.76	-	.40	.32	.24	.28	.15
8.Calmness													
Males													
1965	32	.39	.41	.16	.75	.39	.26	.10	-	.14	-.23	.23	.17
1966	32	.74	.89	.56	.74	.71	.78	.58	-	.06	.21	.20	.28
1967	37	.32	.64	.64	.66	.50	.35	.51	-	.50	.07	-.20	.20
Total	101	.57	.72	.50	.70	.54	.44	.44	-	.28	.00	-.01	.23
Females													
1965	22	.59	.70	.74	.91	.78	.73	.61	-	.32	-.22	.36	.32
1966	38	.38	.50	.45	-.02	.44	.55	.35	-	-.26	.16	.37	.35
1967	24	.28	.52	.55	.61	.37	.50	.37	-	.39	.18	.40	.56
Total	84	.46	.58	.56	.42	.50	.52	.40	-	.10	.10	.31	.38
9.Appearance													
Males													
1965	32	-.10	-.07	.15	.33	.24	.15	.25	.14	-	-.22	.16	.18
1966	32	.09	.07	.07	.29	.00	.06	.45	.06	-	-.03	-.11	-.03
1967	37	.36	.50	.60	.60	.53	.56	.68	.50	-	.01	.03	.20
Total	101	.18	.26	.31	.43	.34	.30	.52	.28	-	.06	-.09	.06
Females													
1965	22	.65	.48	.36	.42	.62	.57	.42	.32	-	.05	.48	.30
1966	38	-.06	.17	.12	.58	.11	-.02	.00	-.26	-	-.16	-.05	-.02
1967	24	.33	.45	.60	.69	.63	.50	.55	.39	-	.15	.02	.44
Total	84	.24	.34	.30	.59	.44	.33	.32	.10	-	.08	.04	.12

skill, and the lesser correlations betwwen those three scales and the four "social scales" suggests a classroom work orientation factor identifying these three scales as somewhat distinct from the four social behavior scales.

A further intercorrelational cluster, although not as sharp as the preceding one, may be noted for three of the four experimental social behavior scales. In particular, scale 1, social range, yields consistent moderate to high coefficients with scale 2, social relationships, r's of .73, .64, .68, .80, .53, and .44 (median .66), and with scale 3, peer acceptance, r's of .55, .50, .51, .74, .54, and .46 (median .53). These latter two scales (2 and 3) are highly correlated, r's of .65, .70, .77, .84, .68, and .86 (median .73). Scale 4, social invisibility, appears more ambiguously tied to these three prior scales, providing correlations of .54, .81, .32, .80, .34, and .52 (median .53) with scale 2, social relationship, but lower coefficients with scales 2 and 3, social range and peer acceptance, r's of .35, .51, .08, .62, .11, and .45 for scale 2 (median .40 and ____[26]),.48, .44, .80, .15 and .55 for scale 3 (median .48).

As shown in Table 3.2, age appears to have no consistent relationship to teacher ratings. In the 1965 sample, 14 of the 18 correlations were negative, ranging from -.41 to .06, with a median r of -.16. On the other hand, in the 1967 sample, 17 of 18 correlations were positive, r's ranging from -.12 to .50 with a median of .20.

IQ's,as noted in Table 3.2, were moderately associated with teacher ratings. The correlations were highest for females, 9 of 27 being .40 or higher; whereas for males only one correlation obtained this magnitude. The correlations were highest for social range (r's ranging from -.23 to .73), social relationship (r's ranging from .02 to .54), peer acceptance (r's ranging from -.11 to .66), attentiveness (r's from .19 to .45), academic skill (r's from .16 to .45), industry (r's from .14 to .42), and calmness (r's from -.20 to .40). The correlations showed greater variability or decreased in magnitude with scales of social invisibility (r's -.05 to .34) and appearance (r's -.11 to .48).

Correlations between T-BR ratings and pupil scores on the Test of Social Inference were discussed in Chapter II for the combined male sample and for the combined female sample in their relevance to TSI validity. The same correlational differentiation commented upon in Chapter II between TSI scores and ratings on certain social scales and between TSI scores and ratings on the additional scales is also relevant to this discussion of the teacher's ability to rate pupils with respect to one attribute differently from ratings on another. As may be noted from Table 3.2, the correlations between T-BR ratings and TSI scores are highest for the social scales, social range, social relationship and peer acceptance and for the scales attentiveness and calmness.

---

[26] The correlations between scales 3 and 4 for 1965 males were not possible because of non-differential pre-ratings given pupils by one teacher on scale 4.

The TSI:*social range* correlations ranged from .29 to .71. The TSI:*social relationship* correlations ranged from .25 to .38. The TSI:*peer acceptance* correlations ranged from .23 to .55. The TSI:*attentiveness* r's ranged from .22 to .42. The TSI:*calmness* correlations ranged from .17 to .56. The TSI:T-BR correlations decreased in magnitude with the scales *industry*, *academic skill*, *appearance*, and *social invisibility*. The teacher discrimination between pupils rated as socially invisible and those rated as at ease in a range of social situations was an unexpected finding but one which supports the argument of the teacher as a discriminating rater.

Examination of Table 3.2 shows that pupil ratings on *social invisibility* were substantially correlated with all scales except *social range* and that these ratings are not dependent upon IQ, hence the *invisibility* scale appears to fit the Sarason and Gladwin prescription of a needed scale as mentioned in Chapter I:

> What is needed is a battery of scales.../*to*/ predict in some degree... the ability to develop social and occupational skills adequate for social living.../to *be*/ divorced as far as possible from the IQ concept. (1958, p.306)

It would be interesting to learn whether the scale might prove predictive of "the ability to develop social and occupational skills adequate for social living."

A general mention should be made of the considerable intersample variability and interscale coefficients to be noted in Table 3.2. Not infrequently these correlations vary between samples as much as .50 correlational points. Two factors might be singled out as contributing to this variability: one, the fact that the samples of pupils differed with respect to age, location (public school vs. institution) and sex; and second, that different teachers were involved in rating each sample of pupils. This latter factor includes, in addition to differences in setting, the fact that for some teachers a characteristic, such as appearance, for example, is of greater concern or is more closely associated in the teacher's evaluation of industry, academic skill, etc., than would be true in other settings. To expect consistent intersample correlations among the various scales would be to argue a negligible effect of intersample teacher and pupil factors. On the other hand, generalization is hazardous. The earlier correlation summarizations in terms of identifiable clusters of "classwork orientation" and "socialness" are based on prevalence (across samples) of interscale correlations, and are offered only as preliminary interpretations from the obtained data. Without question, interteacher variability needs additional directed study.

(b) The consistency of teacher ratings of their pupils was examined in terms of Pearson product moment correlation coefficients computed between the pre- and posttreatment ratings received by the 185 pupils. The retest periods for these coefficients covered a 10 to 13 week interval.

These correlations between the pre- and posttreatment period ratings may be considered a conservative estimate of the consistency of the teacher's discrimination. Some variability in pupil ratings over the retest period is attributable to training effects. Were all pupils to change the same amount and the teacher rating to reflect this, training effects would not affect the pre- and postrating correlations. But this is too unlikely to expect; some uneven changes, one subject gaining much more than most classmates rated like him prior to the training period, is almost certain to occur with respect to certain of the scale criteria over the treatment period of a month and a half to two months. The correlations obtained between pre- and posttreatment T-BR ratings in the 15 classes involved in the experimental phase are represented in Table 3.3.

TABLE 3.3

Correlations between Pre- and Posttreatment Period Teacher Behavior Ratings on Nine Scales

Class	N	TR 1	TR 2	TR 3	TR 4	TR 5	TR 6	TR 7	TR 8	TR 9	Median
A[a]	11	.73	.74	.74	.66	.70	.82	.69	.81	.58	.73
B[a]	8	.60	.85	.96	.92	.93	.83	.98	.95	.96	.93
C[a]	15	.61	.56	.76	.86	.85	.52	.42	.56	.50	.56
D[a]	8	.77	.87	.52	.48	.59	.83	.80	.99	.62	.80
E[a]	12	.87	.74	.83	.51	.69	.80	.89	.66	.84	.80
F[b]	9	.75	.38	.53	.25	.87	.55	.87	.83	.45	.55
G[b]	9	.76	.55	.56	.38	.87	.69	.60	.68	.83	.68
H[b]	9	.40	.86	.53	.73	.73	.56	.76	.76	.14	.73
I[b]	9	.25	.78	.91	.87	.84	.83	.53	.93	.62	.83
J[c]	17	.71	.71	.73	.86	.80	.78	.53	.87	.82	.78
K[c]	17	.44	.73	.55	.64	.81	.55	.73	.63	.86	.63
L[c]	14	.12	.44	.72	-[d]	.56	.88	.71	.77	.49	.56
M[c]	24	.73	.78	.53	.54	.60	.65	.73	.74	.72	.72
N[c]	12	.92	.76	.64	.69	.57	.69	.59	.61	.72	.69
O[c]	11	.83	.84	.75	.56	.94	.23	.92	.87	.69	.83
Median		.73	.74	.72	.65	.80	.69	.73	.77	.69	

[a] In classes A-E ratings were made weekly for 13 weeks. "Pre" and "Post" r's based upon comparison of the average ratings for the initial 3 weeks with the average for the last two weeks.

[b] In classes F-I, weekly ratings were made in the initial two weeks and in the final two weeks of the experimental program. The correlation compares the averages of the initial ratings with the average of the final ratings.

[c] In classes J-O, one rating was made the week prior to and the week following the experimental period.

[d] Assigned identical pre-rating to all pupils.

As is indicated in the explanatory subscript in Table 3.3, the ratings of pupils in the five classes A through E were actually made once each week over a period of 13 weeks. The ratings began three weeks prior to the eight weeks treatment period and continued two weeks thereafter. The correlations in Table 3.3 for these classes are based upon the average ratings for the initial three weeks in comparison with the average ratings for the last three weeks. In classes F through I, ratings began with the treatment week and terminated with the final treatment week of the experimental program. The correlations in Table 3.3 for these classes are based upon the average ratings for the initial two weeks and the week prior to and the week after the treatment. The median of the pre- and postrating correlations for each class is listed in the last column of the table, the medians for the nine scales across the last row. The overall consistency of rater discrimination is revealed by these moderate to high correlation coefficients. Only 7 of the 135 coefficients were less than .40; more than half of the 134 coefficients were greater than .70; a third were .80 or above.

It might be expected that certain characteristics more than others could be more objectively rated by a teacher; that is, be more dependent on pupil behavior than on teacher belief or bias. As an example, a pupil's appearance, his academic skill, his industriousness and calmness would be expected to be more easily observed in a classroom setting than his social invisibility, his peer acceptance, social relationships, and especially, his ease in a range of social situations. Ratings on these latter scales might, therefore, be expected to be more consistent, less dependent on idiosyncratic pupil change in the sense that such change is not as manifest to the teacher. The data does not support this expectation, however, insofar as these pre- and postrating correlations are concerned. There is little apparent difference as to rating consistency from one scale criterion to the next, generally high median coefficients for the scales ranging between .65 and .80. Although the correlations between pre- and postratings on academic skill is high, as one might expect of a highly visible classroom behavior, the correlations obtained on social range scale were of equal magnitude, r's ranging from .56 to .94. Across-teacher differences in rater consistency were somewhat more pronounced, two of the fifteen teachers having median coefficients in the mid 50's, another five having median correlation coefficients of .80 or above. The range of these median correlations is .55 to .93.

The question as to how well a teacher can assess change in rated behaviors can be less well answered within the scope of this project's data. One would expect that ratings made weekly should be more objective than ratings made after the lapse of a considerable period of time because it would be a difficult task to retain as a rating baseline some memory of ratee behavior against which to make a meaningful reassessment a month and a half to two months later. When the rating task is complicated by the need to recall the baselines of a number of persons over a period of time, it would seem from what is known about human "data processing" (Miller, 1958) that the task would be extremely difficult. When the task is demanded of him, it is probable that the rater uses the pupil in comparison with his whole group at the time of the second rating, rather than the pupil in comparison with his original rating. To the extent that this is so, the interpretation of postratings is a dubious business because changes may have occurred throughout the group. The problem is, of course, accentuated in situations where the changes in ratings assigned individuals vary considerably within the group, as would be indicated by a low correlation between pre- and postratings. A more satisfactory situation from the point of view of concern with difficulty of rater task is where all or most of the group gained or declined, as would be indicated by an average gain or loss in rating scores in conjunction with a high pre-post rating correlation.

The pre-post T-BR correlations as shown by Table 3.3, are reasonably high for most classes. The correlations between pre- and postratings of pupils in Class D, however, were not so high for the *social relationship* and *peer acceptance* scales. Retest coefficients for Class F were not high for the *social invisibility*, *peer acceptance* and *industry* scales. Coefficients for Classes H, I and L were not high for the *social invisibility* scale.

The assessment of change in rated behaviors is complicated by the fact that teachers vary both in the portion of the scales they use and in the range of scale intervals they employ. Although most frequent use was made of the central intervals of the scale, differences were noted. As an illustration of teachers' individualistic use of the scale, 5 of the 11 pupils in Class N received pre-ratings from one upwards ("very poor" upwards). An additional five were rated from two upwards ("poor" upwards). Likewise, 13 of the 24 pupils in Class M were rated from one upwards and an additional nine of the 24 were rated from two upwards. In contrast, only two of nine pupils in Class H received a rating below three ("average"). In four of the 15 classes no pupils received a "very good" pre-rating. In contrast, in Class E, nine of 12 pupils were rated "very good" with respect to some scale criterion. The frequency with which each teacher made use of each scale interval in his pre-ratings is shown in Table 3.4 in which the actual frequency of usage of scale positions has been converted to percent.

Table 3.4 about here

An additional factor confounding the interpretation of T-BR gains became evident when the correlation between pre-ratings and the change in ratings (the amount of rating gain or loss when a pupil's pre-rating is subtracted from his post-rating) was examined. It had been anticipated that either near zero or low positive correlation coefficients would obtain, the former indicating that pupils, regardless of initial rating, were as frequently rated higher by their teacher subsequent to the treatment period and the latter indicating that pupils received ratings in relationship to their initially rated capacity.

TABLE 3.4

Mean Percentage of Rating Scale Intervals Used by Teachers of 15 Classes

	Percent Frequency of Use of Scale Intervals[1]								
Class	1.0	1.3 1.5	1.7 2.0	2.3 2.5	2.7 3.0	3.3 3.5	3.7 4.0	4.3 4.5	4.7 5.0
A	4	3	10	10	22	9	26	3	13
B	4	1	19	11	21	11	33	0	0
C	0	1	4	2	12	17	30	24	10
D	0	0	7	14	21	17	32	3	6
E	2	5	6	6	17	7	24	6	27
F	0	0	11	9	32	26	22	0	0
G	0	0	7	14	31	27	21	0	0
H	0	1	5	4	18	17	41	4	10
I	0	10	18	12	23	31	5	1	0
J	17	0	16	0	14	0	33	0	20
K	10	0	18	0	25	0	35	0	12
L	0	0	21	0	55	0	23	0	1
M	20	0	35	0	26	0	12	0	7
N	12	0	28	0	36	0	22	0	2
O	1	0	20	0	26	0	28	0	25
Mean	4.7	1.4	15.0	5.5	25.3	10.8	25.7	2.7	8.9

[1] The averaging of two and three sets of pre-ratings and postratings resulted in the intermediary fractions of 1.3 or 1.5, of 2.3 or 2.5, etc.

The correlations between pre-ratings and changes on each of the nine rating scales separately for each of the 15 classes are presented in Table 3.5. As may be immediately noted, by far most of the 128 coefficients are negative. Furthermore, of the 98 negative coefficients, two-thirds were greater than .50. Clearly, teachers gave higher ratings at the end of the treatment to pupils they rated lower prior to the training program. Although individual teachers varied considerably with respect to the extent of this negative relationship between pre-rating and gain, with two teachers at one extreme having a median coefficient over the nine scales in excess of -.70 and two teachers at the other extreme having median coefficients of .08 and .09, this negative relationship is shown by Table 3.5 to be maintained across teachers. The generality of this relationship across scales is also indicated by Table 3.5. Only two or three of the 15 correlations in each column are positive, these appearing in a total of five of the 15 classes. Half of the positive correlations are accounted for by two of the teachers.

Although it might appear that increased objectivity would accompany ratings made weekly more than those made once before and once after the treatment, the same magnitude of negative correlations were obtained from the seven classes (A-I) whose teachers made weekly ratings as from the classes whose teachers rated the pupils merely before and after the treatment interval (Classes J-O).

TABLE 3.5

Correlations Between Pre-Ratings and Change in T-BR Ratings

Class	Soc. Invis.	Soc. Relat.	Peer Accept.	Soc. Range	Atten-t'ness	Calm-ness	Indus-try	Acad. Skill	Appear-ance	Median
A	-.60	-.11	-.30	-.37	-.29	-.30	-.04	-.68	.14	-.30
B	-.08	-.83	.28	-.54	.12	-.94	.20	.14	-.69	-.08
C	-.10	.26	-.36	.24	-.50	-.51	-.86	-.78	-.61	-.50
D	-.15	-.34	-.56	-.29	-.79	.14	-.60	-.97	-.91	-.56
E	-.58	-.75	-.74	-.80	-.78	-.59	-.59	-.72	-.48	-.72
F	-.73	-.20	-.44	-.07	-.19	-.64	-.46	-.37	.12	-.37
G	.10	-.69	-.48	-.19	-.60	-.50	-.47	-.29	-.19	-.47
H	-.28	.09	.15	.15	.02	-.22	.29	.29	-.79	.09
I	-.38	-.78	-.36	-.28	-.12	-.47	.16	-.24	-.40	-.36
J	-.27	-.86	-.65	-.59	-.86	-.60	-.48	-.63	-.46	-.59
K	-.80	-.90	-.93	-.81	-.85	-.88	-.72	-.80	-.83	-.83
L	-.44	-.06	-[a]	.60	-.27	.04	-.25	.42	-.52	-.27
M	-.11	-.19	-.33	-.47	-.34	-.51	-.09	-.36	-.51	-.34
N	-.64	-.66	-.54	-.52	-.84	-.45	-.68	-.76	-.61	-.64
O	-.38	-.56	-.61	-.79	-.26	-.80	-.55	-.53	-.38	-.55
Median	-.38	-.56	-.48	-.37	-.34	-.51	-.46	-.53	-.51	-.47

[a] All pupils rated "average" makes correlation indeterminate

To some extent this phenomenon may result from the five point ceiling on the scale, which makes it impossible for pupils initially rated maximally to gain. However, relatively few pupils were initially given a maximal rating, as may be noted from Table 3.3; hence, on the whole, there was "room at the top" of the scale. Still, psychologically, as a rater, some teachers may have been reluctant to rate many or even any of their retardates at the top of a scale.

A more substantial and intriguing explanation for the preponderantly negative relationship between initial ratings and rating gains is to posit a desire on the part of teachers to see initially poorly rated pupils progress. The gain indicated for the initially lowly rated pupils may have been subjectively deflected upwards by the teacher's need to see a consequence to his efforts and it is conceivable that this focusing was set in motion by the initiation of the rating process. It is equally conceivable that the working more closely with the academically and socially lagging pupil could result in greater awareness of his gains to the neglect of the less "problem" pupil who may also have gained.

At any rate, for whatever reasons, the usefulness of the teacher ratings as a measure of behavioral change over the experimental period appeared to be severely limited by the interaction of the teacher idiosyncratic use of the scale and his tendency to increase the postrating of the initially low rated pupils. On the other hand, however, teacher ratings do appear to have meaning in terms of behavioral differentiation at a given time. Hence, pre-rating may usefully be compared with pupil scores on the Test of Social Inference, as in Section F3, Chapter II, as an indication that certain of the rated behavioral characteristics are most closely associated with social cue interpretation.

D. Examination of O-BR Ratings

As referred to in Section A of this Chapter and again in Section B (2), because of project interest in pupil behavior in settings other than the classroom, additional "nonclassroom" ratings were obtained for two small classes of pupils at Hissom Memorial Center, a state institution. These 34 EMR pupils, aged 12 to 20, IQ's 43 to 83, were involved in trial 3, described in Chapter V, of the experimental social cue training curriculum. The ratings were made by six persons: chaplain, nurse, librarian, school secretary, director of cottage life, and recreation direction. These six were deemed well qualified as raters because of their acquaintance with all of the pupils to be rated.

Ratings by the outside classroom behavior raters were subjected to interrater agreement. The six outside raters were randomly separated into two sets of three raters, and the averages of the ratings made by one set of raters were compared with the averages of the ratings made by the other three raters in terms of Pearson product moment correlations corrected by the Spearman Brown formula for double test length. The resulting coefficients are reported as Table 3.6.

TABLE 3.6

Correlations between Mean Ratings by Two Sets of Three "Outside" Raters Using O-BR Scales

Class	Soc. Range	Peer Acc.	Social Invis.	Attentive.	Calmness	Industry	Acad. Skill	Appearance	Honesty
				Pre-Treatment Ratings					
J	.76	.90	.86	.82	.78	.92	.92	.94	.70
K	.86	.90	.93	.93	.95	.91	.95	.90	.90
				Post-Treatment Ratings					
J	.75	.93	.44	.64	.68	.61	.75	.87	.74
K	.86	.92	.93	.93	.94	.91	.89	.94	.92

As can be noted in the table, the interrater correlations were generally high, particularly at the time of the pre-treatment ratings. Agreement between raters was somewhat less at the time of postratings, especially with respect to pupils in Class J. Overall, the picture is one of substantial agreement between random sets of outside raters.

Questions such as those asked about the meaningfulness of the T-BR ratings may likewise be asked of the O-BR ratings: (a) Do ratings made by nonteachers of outside nonclassroom pupil behaviors confirm the teacher ratings? Are pupils similarly rated by different raters in both classroom and nonclassroom settings? (b) How well do the outside raters discriminate between the behavioral characteristics they are asked to rate? (c) How consistent are these nonteacher raters in their ratings? (d) How well can they assess behavioral change?

(a) The extent to which O-BR ratings coincided with ratings made by the classroom teacher was examined in terms of Pearson product moment correlations computed between the mean of the pre- and posttreatment O-BR ratings with pre- and posttreatment ratings made by the teachers. These coefficients are presented in Table 3.7.

TABLE 3.7

Correlations between Teacher Ratings and the Mean Ratings by Six Outside Raters

Classes	N	Soc. Range	Soc. Acc.[a]	Soc. Invis.	Attentive.	Calmness	Industry	Acad. Skill	Appearance	Median
				Pre-Treatment Ratings						
J & K	34	.56	.70	.57	.33	.38	.32	.37	.61	.56
				Post-Treatment Ratings						
J & K	34	.44	.73	.62	.34	.38	.32	.45	.61	.45

[a] The definitions of two highly correlated T-BR scales social relationship and peer acceptance were combined into the O-BR scale definition social acceptance. The correlations here are between T-BR ratings on peer acceptance and O-BR ratings on social acceptance.

As may be noted, the coefficients in the table range between .32 and approximately .70 for both pre- and posttreatment ratings. There is somewhat more agreement between O-BR and the T-BR ratings with respect to the social scales peer (social) acceptance and social invisibility than with respect to the academic scales of industry and academic skill. The agreement with respect to attentiveness is among the lower coefficients. One must conclude from the low to moderate r's that the outside rater rates pupils somewhat differently than the classroom teacher with respect to certain behaviors and attributes.

(b) In examination of outside rater discrimination in use of the rating scales, interscale coefficients were computed between ratings made on the nine O-BR scales. In this computation only pre-treatment ratings were examined. Table 3.8 presents these correlations.

TABLE 3.8

Interscale Correlations Based on Pre-Treatment Outside Rater Ratings of 34 Inst-EMR Pupils on Nine O-BR Scales

Scale No.	1	2	3	4	5	6	7	8	9	CA	TSI	IQ
1 Soc. Range	-	.85	.81	.82	.72	.80	.83	.64	.56	.12	.65	.43
2 Soc. Acc.		-	.85	.82	.74	.79	.90	.52	.56	.13	.56	.34
3 Soc. Invis.			-	.93	.91	.91	.90	.80	.71	.16	.32	.26
4 Attentive.				-	.92	.97	.94	.71	.76	.25	.42	.29
5 Industry					-	.91	.87	.68	.77	.32	.22	.16
6 Acad. Skill						-	.91	.73	.72	.26	.37	.26
7 Calmness							-	.70	.75	.17	.42	.27
8 Appearance								-	.50	-.12	.15	.22
9 Honesty									-	.31	.24	.15

As may be noted from Table 3.8, clusters of highly interrelated scales are found, resembling the clusters noted in the T-BR data in Table 3.2. As in the T-BR data, high correlations were obtained between *attentiveness* and *industry* (r=.92), between *attentiveness* and *academic skill* (r=.97) and between *industry* and *academic skill* (r=.91). As rated by outside raters, more than was true of teacher rating, *calmness* was most highly related to *attentiveness* (r=.94), to *industry* (r=.87) and *academic skill* (r=.91).

As was true of the T-BR data, ratings on *social range* were significantly correlated with *social acceptance* (r=.85). Outsider ratings as to *social invisibility* were more highly correlated with the other social scales than was found true of the teacher ratings. The correlation between ratings made on the *social invisibility* and *social acceptance* scales was .85. The correlation between ratings made on the *social invisibility* and *social range* scales was .81. As was noted in the T-BR data, the high correlations between the *social invisibility* scale with other desirable behavioral characteristics suggests that experimental trial should be made of the value of the scale in predicting postschool or post institutional adjustment.

In answer to the question concerning outside rater discriminating in his use of the scales, the generally high intercorrelations showing little of the range of coefficients obtained from the teacher raters indicate that pupils tended to be ranked by outside raters in much the same way regardless of scale.

(c) An examination of the consistency of ratings made by outside raters was made in terms of product moment correlation coefficients computed between the pre- and posttreatment mean ratings. These coefficients are presented in Table 3.9.

TABLE 3.9

Correlations between Pre- and Postratings by Outside Raters on O-BR Scales

Class	Soc. Range	Soc. Acc.	Social Invis.	Atten-tive.	Calm-ness	Indus-try	Acad. Skill	Appear-ance	Hones-ty
J	.86	.95	.91	.89	.93	.87	.92	.95	.73
K	.86	.91	.90	.90	.91	.92	.93	.82	.92

The coefficients in Table 3.9 are higher than those in Table 3.3 for T-BR ratings. Pupil rank order was consistently maintained by the O-BR average rating (the average of the six raters over the 11 week period.)

(d) In considering the objectivity with which outside raters assessed change in rated behavior after a lapse of time, as was computed for the T-BR ratings, correlations were computed between pre-treatment ratings and change in ratings. The use of the O-BR format did not eliminate the negative correlations between pre-ratings and change in ratings, as is shown by the presentation of these data for classes J and K in Table 3.10. However, the correlations obtained from the O-BR ratings, shown in the top two rows, are generally lower than the median of the correlations between pre-ratings and change on the T-BR scales which is presented, by way of comparison, as the bottom row in Table 3.10.

TABLE 3.10

A Comparison of the Correlations between O-BR Mean Pre-Ratings and Mean Change in Ratings with the Median of the T-BR Correlations between Pre-Ratings and Change

Class	Soc. Invis.	Soc. Acc.	Soc. Range	Atten-tive.	Calm-ness	Indus-try	Acad. Skill	Appear-ance	Hones-ty	Median
J	.07	.32	-.35	-.22	-.05	-.19	.01	-.51	-.27	-.19
K	-.16	.03	-.44	-.02	-.31	.15	.11	-.18	-.01	-.02
Median for Classes A-O	-.38	-.48	-.37	-.34	-.51	-.46	-.53	-.51	[a]	-.46

[a] There is no honesty T-BR scale

In summary, the high consistency of the outside rater in rating pupils differentially within their groups, and the more moderate negative relationship between pre-ratings and change lends support to the expectation that the O-BR format would prove a better rating format (and procedure) over the T-BR format. Although no data by way of evidence can be presented from the current trial of the O-BR format, it seems most reasonable that the preliminary recall procedure in which the rater first brings to memory the population selects representatives of the highest, the lowest, the merely average, the between average and highest and the between average and lowest with respect to each scale criterion in constructing a series of explicit frames of reference for each scale, should provide a yardstick of sufficient range and stability to provide for consistency in the rating of change in behavior over time. In other words, the individual names written into the rating scale frame of reference would remind the rater of the population with whom the present individuals were to be compared. One procedure which would perhaps reduce this problem of maintaining equivalent referent frames, permitting more accurate assessment of behavioral change, might be to tape record or otherwise reproduce the rater's referent frames used by him in his initial recall session. The efficiency of these or other possible "stabilizing" procedures is of course only conjectural at this printing.

E. Summary

In overall evaluation of the meaningfulness of the teacher and outside rater behavior ratings, the writers conclude that the pre-treatment ratings by teachers and outside raters were both usable estimations of the pupils' rated behavioral characteristics. Although the ratings were based upon pupil behavioral performance in different settings, the general congruence of the ratings is demonstrated in Section F in Chapter II, in relation to the TSI scores and IQ's of the pupils. The posttreatment ratings by the outside raters are viewed as similarly useful estimations because of the frame of reference procedures described immediately above. Differences between pre- and posttreatment ratings by the teachers, however, whose procedures did not include the recall and frame of reference procedures, are viewed as less clearly interpretable because of the noted admixtures of idiosyncratic use of rating scale intervals, and other problems noted in sections 1 (b) and (c) earlier in this chapter. Although the outside rater procedures and format (O-BR) as given trial in the present project appear superior to the teacher rating procedures (T-BR), this cannot be demonstrated within the present project.

IV. SOCIAL PERCEPTUAL TRAINING MATERIALS

To develop training and educative procedures to improve the social comprehension and social behavior of the retarded adolescent, a major portion of the project consisted of planning, preparing, trying out, and revising experimental lessons. The exploration of several modes of experimental training and a discussion of this project's particular approach to experimental training will be presented in this chapter, as will be the project's decision to prepare a more elaborate educative program than might be envisioned as remedial for only a deficit in social cue interpretation. In this connection, the social anthropological perspective that shaped the project's educative program is described. Further, some of the limitations in the young retardate's behavioral repertoire are discussed as determined from interviews with local agency personnel concerned with post-school adjustment of the retarded. From these interviews with training supervisors, counsellors, and operators of boarding homes came ideas for the content of the experimental lessons. A section describes the general format of the materials and the contents of the four sets of lessons that were developed sequentially, making use of classroom trials. Teachers' criticisms of the materials, consultant suggestions, and information from the agency interviews were the basis for revisions.

A. Social Cue Training Alternatives

Several procedures which might be expected to improve the retardates' social situation comprehension were considered for experimental trial. One approach would involve the preparation of illustrations of social cues--pictures of facial expressions, gestures, posture, or clothing, etc., on a continuum of least ambiguous to most complex for use in discrimination training. It was reasoned, however, that it would be difficult to involve the classroom teacher or the supervisor of training in following this laborious process for which there was no previous experience in the special education classroom agenda. As a second general type of approach, the remedial literature contains studies in which discriminations and responses have apparently developed in less unitary a fashion when subjects have witnessed an instructional film or enactment of behavior. (Bandura & Walters, 1967, pp. 49ff)

As precedent, Kraft (1962) used excerpts from five movies depicting young adults in realistic situations in developing a test to measure the abilities of young adults to understand others. Goodman (1962) used recordings from plays to accompany an instructional handbook for teachers to use in increasing understanding of social interaction. The attempt to produce its own filmed material would, however, have greatly exceeded the present project's resources. Although this lead was not followed, there is evidence of the value of movies in the education and training of the retarded (Driscoll, 1962) and it is speculated that such an approach could be an effective one.

As a third possibility and the approach that was finally chosen, there remained the use of filmstrips or slides in conjunction with tape recorder tapes, as a means of social cue interpretation training. By preparing a teaching guide or manual and accompanying illustrative materials, it seemed possible to arrange a series of lessons that could be easily used in a variety of instructional settings and be easily supplemented by an instructor in situations if additional explanation or emphasis were needed. The use of slides and tape recorder tapes suggested the possibility of looking beyond the immediate goal of improved cue interpretations toward the more broadly functional goal of explaining the patterns of adult activities. Fudell (1963) in his prevocational lessons for retarded youth makes the point, very strongly, that what is expected of the youth in school settings and what is expected of him as a worker in post-academic life, are widely divergent. Because of the "distance" that separates the activity world of the average adolescent from the activities of the adult world, needed are illustrative samples of functionally important social settings and activities to stimulate classroom discussion and examination of their identifying characteristics. In the education of the nonretarded pupil, geographically distant places are illustrated to acquaint pupils with the appearances and practices of other cultures. In a similar way it was planned to illustrate some of the commonly used places or settings in the "outer world" of the grownup as the foci of discussions of the what or how and why of participant behaviors.

B. The Social Anthropological Outlook

Norms: The manner of dealing with that "outer world" was influenced by the writings of anthropologists and social ecologists.Goffman (1959, 1961,1963), Hall (1959), Barker & Wright (1955) and Barker & Gump (1964) were helpful in illustrating the universality of norms as guides to behavior. These writers reveal by their examples that much interactional behavior is predictably patterned through conformity to cultural conventions. When one thinks of activities as a series of somewhat ritualized performances taking place in identifiable settings and subsettings, one has the means of breaking the behavioral flow into manageable portions. There are approaches, transactions and leave-takings. Behavior is modulated by the event in which one is participating or spectating. There are highly predictable sequences of behaviors that "go with" situations or places. In going to church or in taking part in a wedding or a funeral, for example, it is probable as a preliminary that one has made changes in dress and grooming, and that participant behaviors will resemble in certain ways the behaviors of others present. It is probable that clothing will be of more subdued color, of better quality, in better repair than clothing that would be worn for transactions in the supermarket. Grooming would probably have included attention to fingernails, hair, and shoes, and would be more apt to include bathing, shaving, use of deodorants, and changing of underclothing that would a trip to the supermarket. Why this difference in preparation? In a sense, it is preparation for admission. One is accepted when one's appearance "fits" the expected. One may not be accepted if appearance and behavior deviates from the usual range. Deportment or performance is similarly subject to modulation to the convention. In the foyer of the church or funeral parlor, peripheral to the main setting and main event of worship, wedding, or funeral, the participant is apt to look about for persons he knows, greet them, and speak briefly with them. At church and at the wedding, but not at the funeral, it is probable that participants to the encounter would smile. During the main event participant or spectator behavior is restrained in comparison with behavior at less formal situations, as, for example, at a beach party or a baseball game.[26]

Behavior Settings: Barker & Wright (1955) and Barker & Gump (1964) speak of behavioral patterns as being generated by the action setting. Although the obverse is true; namely that the setting has been constructed by humans to facilitate their performances, it is also apparent that the pattern of behavior is predictably modified by identification of the setting. As an example of a behavior setting, Barker mentions a school classroom. There is a boundary surrounding the classroom. Inside, the activities and physical objects are not random. There are predictable types of physical objects; e.g., chairs, chalkboards, desks, cupboards, etc. The patterns of interaction between persons, and between persons and physical objects are discriminable from those that occur in other settings. (Barker & Gump, 1964, pp 16-17) There is predictability to the pupils' performance because it is dependent upon a teacher's direction and the teacher conducts the scene in accordance with the cultural expectancies for classroom conduct.

Admission Credentials: To gain access to behavior settings there are requirements for admission. This allows some control at the point of entry permitting the rejection of those individuals anticipated as being disruptors or otherwise non desired. But even after admission, there is the requirement of continuing to measure up to transactional standards, and those who do not are ejected. This phenomenon of admission can be recognized when one thinks of private clubs, fraternities, institutions of higher learning, and so on, and the "measuring up" requirement recognized when one thinks of probationary periods. The extent of these phenomena--their universality--must also be emphasized. To become aware of the universality of these phenomena, consider for the moment the world of man's construction. Everywhere there are areas enclosed by walls to which there is entry via valve-like gates and doors. Some of these areas are "private" places whose admission is carefully restricted to persons known to the owners, or to officials whose status can be known from their apparel or from presented credentials. Although access to "public" places is less restricted, it is common for public places to have conditions for entry as for example:

---

[26] There are wide differences in subcultural conventions, and the formal behaviors described above are relevant to a particular subculture. In other subcultures a funeral might be celebrated with a brass band; a wake might be a convivial occasion. However, in order to fit in with others, one must know what is expected and when, or must be observant and able to modulate behavior imitatively. Although less restrained behavior, (loud singing, shouting and clapping) as an example, may be appropriate in some churches, even this behavior must conform to an expected course or pattern in that setting. One must sing the prescribed songs; shout the expected utterances; and clap at the expected times. To do otherwise is to be deviant.

Hours 9:00 a.m. - 6 p.m.
Admission $1.50
Children under the age of 14 are not admitted to patient's rooms
Beach apparel is not allowed in the dining room
Employees only
Men
Women

The vocational preparation of the retarded customarily calls attention to a single such situation--the job screening interview. The interview is to determine whether one will fit. One's general appearance and deportment are presented with written credentials, if possible, as evidence of admissibility.

Aside from the requirements for acceptance into physical structures, there are also friendship requirements. Acquaintances are permitted closeness and are invited to extended interaction depending upon the "credentials" that are detected or inferred from general deportment. In this regard, interactional signals bounce rapidly back and forth, perhaps akin to the porpoise's sonar. In some way, signals are decoded, at least by the nonretarded, as indicating admission or nonadmission, and as further indicating the level or degree of admissibility.

Viewing the adult's milieux from this perspective, it appeared that the project's experimental lessons should attempt to illustrate the outer world as consisting of places that one may enter at appropriate times, in conformance with either explicit or implicit behavioral requirements which may be recognized through visual cues. As an example, a sample lesson is included as Appendix D . This lesson entitled "Places Sometimes Tell Us What to Do" is the first lesson in the third week of the final version of the curriculum.

C. The Retardate's Behavioral Repertoire[27]

The next concern of the project was to learn more about the retardate's maladaptive behaviors--the "credentials" that he lacks--that keep him on the outside. Toward this end, (1) the literature of the retardate's post-school social adjustment was reviewed in a search for examples of behavioral specifics; (2) direct observations were made of sample populations of retarded in state institutions and community settings; (3) information was sought from the agencies directly involved in vocational preparation or other phases of post-school adjustment.

The Literature: The literature yielded many references to the social exclusion or difficulties of the retarded, but relatively few references to the specifics of his behaviors. However, Baldwin (1958) listed reasons given by nonretarded peers for their dislike of their retarded classmates: "He talks back to the teacher"; "He fights too much"; "She can't do anything"; "He talks too much"; "He never studies"; "He disturbs our class"; "He can't play." Johnson (1950) listed similar reasons given by nonretarded peers for the dislike of their retarded classmates: "Misbehaves," "Poor sport," "Cheats," "Dirty," "Smells," "Poor in school work," "Copies," "Different," "Too big," "Peculiar speech," "Shows off," "Tries to attract attention," "Steals," "Swears," "Calls names," "Lies," "Gets others into trouble," "Rough," "Mean, bullies, teases, fights," etc.

---

[27] Partial support for this portion of project activities was a research fellowship grant to the University of Kansas Medical School from the National Association for Retarded Children. Appreciation is due Miss Elizabeth Galton, fellowship grantee, for her ideas and assistance.

Direct Observation: In addition to the search for behavioral references in the literature, behavioral observations were conducted in public school special classrooms, in three state institutions for the retarded in Kansas, and at several social events arranged for junior high school aged retarded pupils as a part of a local interschool "coed" club. In these observations, the general behavioral deviance noted was the failure of some groups and of some individuals within other groups to behave as though they were aware of the usual conventions for a "performance."

During the classroom observational sessions, it was the impression of observers that classroom behavior did not appear deviant in situations in which teachers transmitted an agenda or program and "stage directions" to their pupils. The day's agenda would be posted and discussed so that pupils were aware of the programmatic and behavioral expectancies. As example from a demonstration class which was in daily session for six hours: "It is time now for us to talk about our program for the day. While we talk about our program, face me and watch what I am doing." Toward the conclusion of the scheduled activities, class attention would be called to the next item on the agenda so that transitional behaviors would be less random: "What is next after reading, Ken?" and a discussion of how to make a smooth transition: "You can leave your pencils and notebooks out but put the other things away." In the same demonstration class pupils were encouraged by the teacher to think evaluatively about themselves and one another.

"Joe, listen to your voice when you talk; see if you can tune it down."

"How does Tom look this morning with his new slacks and clean shoes?"

The same teacher made use of a pupil self rating form on which pupils gave themselves a rating at the end of the day with respect to behavior the pupils themselves selected as being important. A second demonstration teacher, in addition to using the above techniques, combined the use of bonus stamps contingent upon the recording of improved academic and social performance. (Stables, 1966)

In contrast with the generally appropriate classroom behaviors noted in the demonstration classes, in other situations in which the agenda, the "coaching," and the contingencies were lacking, many disruptive behaviors were observed. The importance of the class schedule, the teacher's explicit directions, and quick application of "consequences" in maintaining conventional behavior was demonstrated when student teachers were assigned to a demonstration class. One novice teacher's failure to use a posted agenda and to tell and remind pupils what was behaviorally expected was accompanied by the appearance of work-disruptive pupil behaviors. Although some pupils behaved as though they had acquired the conventional schemata, others appeared to need the directions and the contingencies.

Behavioral observations at the state institutions also revealed that while the appearance and visible behavior of many residents fit customary norms, other residents were highly conspicuous because of their dress and mannerisms. Some individuals shuffled rather than walked. Many had conspicuously poor posture. In some institutions residents might be dressed as though they were younger than their size would suggest them to be. Some individuals would approach a visitor too closely. Directly addressing him or touching his face, hand, or clothing, they might repetitously ask his name, where he was from, etc. In one setting this behavior was apparently more than tolerated by the attendants and nursing personnel who behaved toward the residents as though they were darling children. In that setting there were numerous attention-getting confrontations of visitors by the adolescent "children." In contrast, retarded residents of similar age and IQ levels in other settings dressed and behaved appropriately in accordance with their age.

An observer visit to a junior high school coeducational social club produced no record of deviant behavior. This social club had been established by a group of nonretarded high school youth, the Teen Age Monitors, interested in learning to work helpfully with the handicapped. Monitors would plan a social activity once a month during the school summer vacation and would be assisted in the plans by some of the retarded young people. The Monitors who attended the parties viewed themselves as providing behavioral models. The project observation occurred at a theatre party at an outdoor theatre. The retarded boys and girls were not noted as behaviorally different. The club sponsor commented that the situation had been different when the group had gone to a movie during the winter as the boys did not know what to do about coats.

Information from agencies: As an added source of behavioral information, interviews were held with representatives of agencies involved in the post-school adjustment of retarded youth. These conferences were productive of information concerning the retardate's unpreparedness for general community acceptance. A resume of examples of conspicuous or deviant behavior may be noted at the conclusion of this subsection (c). Information was acquired from the following agencies: [28]

The Kansas Work Study Demonstration Project, Kansas City, Kansas
Vocational Rehabilitation Unit, Topeka, Kansas
Goodwill Industries, Kansas City, Missouri
Parsons State Hospital and Training Center, Parsons, Kansas
Paseo Manor Boarding Home, Kansas City, Missouri
Watson's Boarding Home, Kansas City, Missouri

---

[28] The Kansas Work-Study Demonstration Project was involved with the vocational readiness of retarded youth. In its program, retarded pupils at junior and senior high school levels were placed on a series of jobs in actual (nonschool) work situations a portion of each school day. The cooperating employers reported both the adaptive and the nonadaptive work skills of these trainees to the (cont.)

The problems of the retarded cited by representatives of these agencies and homes rarely included failure to pass a job screening interview. It appeared, in other words, that vocational training and counselling had been generally successful in shaping the trainee's behavior so that it was adaptive within the interview setting. However, there were numerous references to maladaptive behavior which made trainees conspicuous in job settings. Examples of these, but not a comprehensive list, are the following:

Wears clothes that are too nice for the job
Refuses to do the required "dirty" work
Doesn't know how to put on an apron and tie the bow
Wears new hat to work and refuses to take it off
Talks too much with other employees and with customers
Talks too loud
Tells tall tales that irritate employees

---

28 (cont.)
vocational counsellors employed by the Work-Study project. The project maintained for each pupil a file of employer and counsellor reports focusing upon his behavioral adequacy. These reports, in addition to conferences with the project staff, provided many instances of behavioral inadequacy.

The vocational Rehabilitation Unit in Topeka was involved with intensive vocational and life-adjustment training of groups of retarded, usually groups of those who were in transition from institutional residence to community residence and employment. The staff social worker maintained contact with many of the former trainees and was extremely knowledgeable as to their problems.

Goodwill Industries had had much experience in training retarded persons for placement in competitive employment. In addition to their awareness of maladaptive behaviors which interfered with employment, the training supervisors at Goodwill were in close touch with several boarding facilities housing a number of their employees, and were aware of some of the difficulties in those settings.

Parsons State Hospital and Training Center at Parsons, Kansas, a state institution for the educable retarded, has been successful in training a good percentage of its residents to resume roles in the outside community. Members of the institution's staff were knowledgeable about the problems which at times led to the failure of residents to make a successful transition to independent living.

The Paseo Manor and the Watson Boarding Homes were privately operated residences in Kansas City, Missouri, housing, as boarders, a number of transitional individuals from state institutions for the emotionally disturbed and the retarded.

Approaches clients or customers too closely, speaks to them too familiarly
Asks questions all the time
Is late to work because won't use public transportation
Is absent without notification
Quit without giving employer notice
Doesn't listen to or remember verbal instructions
Fails to go promptly to his work station--fools around
Fails to remain at his work station
Fails to carry through on assigned tasks
Works too slowly
Can't stack canned goods
Can't use a paring knife without getting cut
Is awkward and slow using tongs to lift fried potatoes into sacks
Can't make change
Can't write down customer's names
Doesn't see any reason to work
Thinks the job isn't good enough
Thinks he should have more money
Gets upset when work is criticized

Another list of behaviors made the trainees appear deviant in settings other than the vocational. Among them were the following:

Doesn't know when clothes are the wrong size[29]
Doesn't store clothes adequately (doesn't know use of chest of drawers and hangers)
Doesn't recognize when clothing is soiled or unpressed
Doesn't know what to do with soiled or unpressed clothing
Doesn't know how to make simple repairs to clothing
Doesn't know when clothes look right for an occasion
Gets into argument with roommate
Lends and loses his belongings
Feels upset because of having no place to go on vacation
Has no idea what to pack in a suitcase when he goes home to visit
Won't bathe
Doesn't use deodorant
Gets ingrown toenails because don't know about clipping them
Doesn't know how to use a telephone
Doesn't understand money
Writes checks as long as he has any checks left
Sells his belongings (or his medicines)
The stores don't like the girls to come in because they just stand around and handle things

---

[29] Examples were given by Goodwill training supervisors of a boy who walked to work everyday in shoes that were much too small. His feet became blistered and bloody before Goodwill personnel found the problem. As another example, one of the girl trainees thought she was "lovesick" because she hurt. She hurt less when it was discovered she was wearing a size 32 bra when her actual size was 34.

Arrested for taking lunch meat from supermarket without payment
Disliked by police because talked funny
Some don't know what it means to "be friendly"
Some girls don't know anything about their physiology and hygiene
Some girls don't know what "dating" means
Goes off with any man who is nice to her

Other behaviors, although not unique to the retarded, were viewed as maladaptive:

Got involved in time payments
Got signed up for expensive insurance
Got signed up for expensive health club membership
Spends his time at the girlie movies and in the bars
Arrested as prostitute
Beaten up by "hoods" on the street
Picked up on the streets by cruising homosexuals

D. Social Perceptual Training Lessons

The decision to prepare visual and auditory cue training lessons for the retarded adolescent posed questions of lesson format and content that are discussed in this section. In all, three lesson series (Curricula I, II and III), differing to some extent in content and presentation, were given trial. After a brief explanation of the lesson format, each of the three experimental curricula, as well as the fourth and final version of the materials (Curriculum IV), will be described in this section, followed by a discussion of the motivational devices employed in the lessons.

Format: Most classroom curricular material is prepared in outline form describing the lesson objective and listing books and resource materials related to the development of the topic. The teacher is usually familiar with the stated educational objective, and has prepared and acquired a store of related exercises and illustrations by which to convey information and develop responses. In this instance in which experimental lessons were to be assigned to teachers on a random basis, the illustrative materials and even the objective would be new to them. Although the objective of social behavioral competence accords with the goal of special education social studies programs, the task of cultivating attentiveness to visual cues for the identification of types of locales and associated behaviors was expected to be somewhat new. It appeared that the experimental lessons would have to be completely written out and illustrated if the factor of differences between teachers in the presentation of new material were not to obscure any possible effectiveness of the material.

The project task, therefore, was to write a teaching script for a series of daily lessons. A general topic or unit was presented throughout several lessons. Each lesson included a statement of the lesson's main objective; also included was a list of the equipment, such as tape recorder and slide projector, that would be required for the day's lesson. In addition a list of vocabulary words that would occur in the lesson and which might require teacher explanation and clarification was presented. As an example of the general format, one of the lessons is included as Appendix D . As can be noted from the sample lesson, the teacher made an introductory statement, then used the first slide. The slide illustrated the point of discussion and was used to elicit pupil response. Lessons were prepared for verbal presentation without requirement that the pupils read or write.[30] As a part of the format, a verbal quiz would follow every several lessons. Pupil verbal responses to the quiz would be tallied as correct or incorrect. All of the illustrative materials: slides, tape recorder tapes, dittoed seatwork, driver handbooks, play money, merchandise cards, tooth paste cartons, etc., and miscellaneous articles such as tape measures, playing cards, and mail order catalogs required for use with the lessons, were provided in kit form.

Lesson Contents: Earlier, in Section A of this chapter, the general objective was mentioned of using illustrations in the classroom of often-frequented adult settings as foci of discussions concerning the what and why of participant behaviors. This objective led to the preparation, over a three-year period, of more than a hundred lessons illustrated by photographic slides, recorder tapes and other items that were tried out in sets of from 40 to 50 lessons to a program series.

The first set of lessons (Curriculum I) focused most directly upon cue interpretation, and to a lesser extent upon the difficulties of community adjustment noted by the agencies mentioned earlier in this chapter. The second set of lessons (Curriculum II), in part a revision of Curriculum I, included new contents bearing upon some of the retardate's problem areas as described by the agencies. As an example, frequent reports of trainee dissatisfaction with certain occupations, viewed as low status, led to preparation of several lessons on restaurant work and work in a large laundry. It was reasoned that familiarity, even via pictures, might tip the scales somewhat in favor of the more familiar occupations. Reports of the difficulty of finding wholesome inexpensive recreation led to material of that nature in Curriculum II, including practice at card games. The third set of lessons (Curriculum III) diverged even more from simple cue interpretation. Its lessons ranged more broadly over the retardate's problem areas, including money management, shopping behaviors, choices of residential facilities, and how to behave on a date.

The three curricula that received actual classroom trial and the fourth curriculum representing the culmination of trial and revision are briefly described as follows:

Curriculum I: Curriculum I consisted of a teaching script (manual) for 40 lessons. Two hundred fifty-three 2x2 slides and two tape recorder tapes illustrative of visual and auditory cues discussed in the lessons accompanied the manual. Also supplied for use with particular lessons were several dittoed forms and a plywood model city layout, with toy cars and buildings. The topic headings for Curriculum I lessons are suggestive of the lesson contents:

---

[30] The project directors felt it desirable that pupils learn it was possible to function in a community even without the ability to read. The shift from the traditional academic emphasis on reading and writing was highly motivating to certain pupils. (See teacher comments on page 5.18, Chapter V. This was confirmed by Clark (1967b)).

Week I	Introduction to Signals
	Hand and Arm Signals (Gestures)
	Traffic Signs and Signals
II	Places and Things Sometimes Tell Us What To Do
	What To Do and What To Say
III	Field Trip
IV	Signal Properties of Clothing
	Reading People: Reduced Cues
	Knowing People by What They Do
V	Where We Spend Our Time
	Where and With Whom We Spend Our Time
VI	People Having A Good Time
VII	Best Social Behavior: Being Grown Up and Proud
VIII	Review & Test

Lessons of the first week of the series called attention to signals that tell people what to do. The material began with signals that pupils were familiar with, such as the upraised hand of a policeman to stop traffic, the flashing red light and siren of a fire truck; then discussed signals expected to be somewhat less familiar. In the second week the lessons pointed out that many things around us resemble signals since their implications can be read. (Ruesch, 1956) As an example, fences and padlocks are a way of saying "Keep Out." We can sometimes "read" how we are supposed to behave toward machines by looking at the way they are made. A coin slot, a lever or knob are cues as to behavior; as examples, vending machines were illustrated. We can sometime see what is expected of us by looking at the structure of a setting. As examples, doorways, aisles, sales counters, checkout stands were illustrated as behavioral cues.

In the third week a field trip to a laundromat was prescribed so that pupils could have a demonstration of the use of a series of coin operated machines. The trip was prefaced by a series of slides illustrating the use of the machines.

In the fourth week, lessons pointed out that because vocational clothing is functional, one can infer some things about people by looking at what they are wearing. Protective clothing is worn by construction workers, umpires, football players, astronauts; warm clothing is worn by people who work out of doors; showy uniforms are often means of identifying an employee of a particular organization; washable uniforms often mean that the wearers are employed on jobs that may get them dirty, as by mechanics, waitresses, kitchen workers, beauty operators, surgeons, etc.

The last three weeks of the program attempted to focus the retarded pupil's attention upon persons he associated with and how they felt, as might be indicated by visible and auditory cues. This portion of the curriculum also discussed the pupil's own feelings as consequences of his behavior and, in Weeks VI and VII, indicated that pleasure grows from learning how to do things better, from learning the rules, from behaving in a more grown up way, and from learning appropriate things to say and do. In Week VII a discussion of what to say in various situations was illustrated by slides and accompanied by role playing.

The final week of the program consisted of quizzes in which questions were asked about slides of social scenes. Correct responses were entered on a chart; when sufficient areas on the chart were filled in, the pupil would earn a nickel.

Curriculum II: The teaching script for Curriculum II consisted of 48 lessons illustrated by 436 slides and a tape recorder tape. Seatwork materials were primarily self scoring forms for quizzes. Sundry additional illustrative and practice materials were provided as accompaniment to particular lessons. Among these were driver handbooks, sets of cards illustrating merchandise items, sets of play currency, toothpaste cartons, and bank deposit slips.

The lesson contents are suggested by the following topic headings:

Weeks I - III (virtually identical to Curriculum I)
IV Transactions: Several Kinds of Stores
" Where We Go for What We Need
" Entering, Specifying, Inquiring, Deciding
" Entering, Trying, Fit, Appearance, Decision to Buy
V Reading People (virtually same as Week IV in Curriculum I)
VI Living On Our Own
Finding a Place to Stay
Looking Clean and Neat
Living on our Own: Some Problems
VII Jobs and What They Take: Money Management
" " " " Being on Time
" " " " Being Careful
" " " " Field Trip
VIII Listening and Following Directions
What It Takes to Get and Keep a Job
Getting the Picture of What's Going On
Getting Along with Others
IX People Signal How They Feel
What We Say Makes a Difference in the Way People Feel
Getting Along with Others
X Review

The first three weeks of Curriculum II closely resembled the initial three weeks of Curriculum I. The fifth week closely resembled the fourth week of Curriculum I. New Material and lessons were prepared for Weeks IV, VI, VII and VIII. In the fourth week, slides illustrating the exteriors and display windows of several types of stores were presented as a focus for discussion of merchandise categories; pictures of various commodities, pasted upon cardstock, were supplied with the kit for pupil sorting practice. In the same week there was discussion of appropriate shopping behaviors--entering, specifying needs, and inquiring about items; there was also some role playing practice.

The sixth week began a story in which a young man, Fred, travelled by bus to a city to find a job. The bus trip story was illustrated by slides showing the central figure at the ticket counter, storing his suitcase at a baggage locker, sitting on a waiting bench, standing in line to enter the bus, giving his suitcase to the driver for storage, presenting his ticket; then, later in the story, finding the YMCA, walking to the desk clerk's desk, finding the room number that matched his key tag, unpacking his suitcase, grooming himself prior to going down to the lobby, and so on. In the course of the week's lessons, Fred found a decent rooming house, visited a hospital waiting room, a dental clinic, etc.

Play money was furnished for use with lessons in the seventh week in which the pupils, emulating Fred, were to "pay" their weekly board and room, set aside their bus fare, save money for a hair cut, and see how much they had left from a small weekly salary.

In the eighth week a chart illustrated behavioral stages that are requisite to employment. Also illustrated were such requirements for retaining a job as following directions, being aware of what others are doing, getting along with other workers, and so on.

In a ninth week some Curriculum I illustrations of interpersonal cues supplemented by new materials were used in the lessons, "People Signal How They Feel." A tenth week was a week of review.

Teacher evaluations of this curriculum were based upon its use in institutional settings with pupils ranging in age from 11 to 20, having IQ's between 41 and 84, in comparison with a second set of experimental prevocational lessons[31] as is described in Trials 2 and 3 of Chapter V.

---

[31] Fudell's (1963) Occupational Educational Units (recently published under the title How To Hold Your Job (Fudell, 1967) consisted of prevocational lessons intended to instill values and attitudes basic to vocational success, developed for a 12 week program. The daily lessons in this curriculum, akin to the social perceptual lessons, were presented in the form of verbal interaction between teacher and pupils. The teacher's guide, although not prepared in verbatim form, carefully delineated the teacher's role; the opening and closing statements being written out verbatim. Handout sheets were furnished as a basis for pupil reading; words were carefully defined; pupils kept notebooks of new vocabulary. A weekly quiz, unlike the social perceptual lessons, required pupils to respond in writing to written questions prepared at second grade reading level. Aside from condensing the lessons to make them coincide with the 10 week experimental periods, lessons were presented without modification. A list of the lesson headings is presented as Appendix E . In their trial at Parsons they were supplemented by filmstrips selected as relating to the lesson topics. A list of the filmstrips and movies used to accompany them is presented as Appendix F .

Curriculum III was an extensive revision of the prior materials. The teaching manual contained 50 lessons, none of which were identical to lessons in the earlier versions of the curricula. The illustrative materials were similarly extensively revised. A major portion of the 440 slides were newly prepared. An accompanying tape recorder tape was newly prepared. Seatwork materials to accompany several lessons were provided teachers in the form of Dittomasters. Additional illustrative materials furnished the experimental teachers included State driver handbooks, mail order catalogs, bank deposit slips, merchandise cards, tape measures, rules for card games, and decks of playing cards.

The following list of topic headings is presented as suggestive of the lesson contents:

Week I	Introduction to Signals
II	Using Numbers as Signals
III	Places Sometimes Tell Us What To Do
IV	Making a Good Impression
V	Where To Go for What We Need
VI	A Department Store
VII	Living on Our Own
VIII	Living at Home
IX	Having a Good Time Because of Knowing How
X	Review

The introductory sections on signals were redrafted at a higher level of language and conceptual content. In consideration of certain of the agency-cited problems of the retarded, new lessons illustrated the system of street markings and use of public transportation. New lessons were added on neatness and cleanliness and on keeping possessions safe and in good repair. Lessons on shopping were expanded and were followed by a field trip to a department store where both consumer and employee behaviors might be noted. The narrative about Fred was elaborated to illustrate efforts to cope with a barely adequate income. A high school girl, Linda, added as a second narrative figure, was the focus of a new set of lessons providing examples of some adaptive on-job behaviors. Fred's date with Linda was illustrated and discussed.

Curriculum IV: In the final preparation of the social cue training lessons (Edmonson, Leach & Leland, 1967) lessons from the first eight weeks of Curriculum III were retained with minor revisions. The Curriculum IV ninth week lessons concerning leisure time activities were expanded by the addition of lessons from Curriculum I into two full weeks of material. Curriculum IV consists of 50 lessons illustrated by 440 slides, tape recorder tapes, and dittoed seatwork materials.

Motivation: As a device to assure pupil motivation, the lessons in each curriculum prescribed rewards and competition for correct responses to several quizzes. In Curriculum I reward and competition were prescribed in six lessons. Consequent upon his correct responses a pupil "moved up" a ladder chart a rung at a time and, on attaining the top, received a nickel from the teacher. As added incentive, in the first three quizzes, the teacher assigned pupils to a team. Each member of the winning team received an extra credit in addition to his individual earnings. Team assignment shifted from quiz to quiz. In an effort to "shape" more comprehensive pupil responses to the quiz pictures, in connection with two quizzes credits were given only for comprehensive replies. In the trial of the curriculum (described as Trial 1 in Chapter V) rolls of nickels were supplied the experimental teachers.

In Curriculum II similar procedures of reward and competition were written into ten lessons. The reward ladder and a similar pathway chart on which the progress was rewarded by nickels were used in connection with several quizzes. To encourage complete responses, in five quizzes, credits were given only for comprehensive replies. Team competition was used in connection with one quiz. In one trial of the curriculum (described as Trial 2 in Chapter V) the project paid for canteen coupons used by the experimental teachers in lieu of cash.

In Curriculum III, in lieu of nickels as a reward, a savings point system was introduced. Correct responses to quizzes in 20 lessons resulted in the accrual of points in savings books. During the trials of Curriculum III pupils could exchange their accumulated points for items of consigned merchandise. The initial 15 quizzes concerned comprehension and recall of material discussed in the lessons. The last five quizzes concerned interpretation of visual materials. Team competition was used in two quizzes.

In the trial of Curriculum III (described as Trial 4 in Chapter V) boxes of merchandise items were consigned to each experimental class to be exchangeable for accrued points. The merchandise assortments included pocket combs, nail clippers, cosmetic cases, jewelry, head kerchiefs, men's handkerchiefs, kleenex, socks, billfolds, pens, pencils, notebook paper, erasers, flashlights, batteries, and other items.

In Curriculum IV the importance of extrinsic motivation to the performance of retarded subjects is discussed and the savings point system described, leaving the system of exchange to the individual teacher.

E. Evaluation of Lessons

Four trials of the social perceptual curricula conducted in nine classrooms during the experimental phase of the project are described in detail in Chapter V. The trials were carried out under conditions far from optimizing the effectiveness of either the social perceptual training lessons or the materials used by way of comparison. To particularize, teachers had no choice in their assignment of materials and they received the materials barely ahead of their use. Their orientation to the social perceptual materials was a written statement of curriculum purpose provided as an introduction to the lessons. Their orientation to the filmstrips and movies used as contrast materials in Trial 1 was merely a request for an evaluation of their suitability for institutional pupils. Their orientation to the contrast material in Trials 2 and 3 was the written statement of curriculum purpose provided by its author as its introduction. The teachers had a limited span of time in which to cover lessons with which they were unfamiliar. They were not able to engage in long range planning with frequent repetitions of material to the point of thorough assimilation.

Despite all of these difficulties accompanying the experimental trials, significant gains in social cue interpretation were found associated with use of the social perceptual lessons in all four trials and with use of the contrast materials in Trials 2 and 3. Use of the social perceptual Curricula I and II in Trials 1 through 3 resulted in greater gains by boys than by girls as measured by the Test of Social Inference score. Use of the contrast curriculum in Trials 2 and 3 resulted in greater TSI gains by girls than by boys. Use of the social perceptual curriculum III in Trial 4 resulted in equivalent TSI gains by both boys and girls.

Use of the experimental Curriculum I in comparison with audio-visual supplemental materials in Trial 1 resulted in somewhat more frequent gains on four of five social behavior rating scales by initially highly rated pupils in the experimental classes than by pupils of similar initial ratings in the "placebo" classes.

Use of experimental curriculum II in comparison with the contrast lessons in Trials 2 and 3 did not show a similar difference in behavior ratings favoring the social perceptual lessons.

Use of experimental Curriculum III in comparison with a "no treatment" control class in Trial 4 resulted in no significant differences in behavior rating.

Despite teacher unfamiliarity with the social perceptual materials, the occasional errors that were found in this material, and the limitations upon teacher effectiveness that were described earlier, teachers reported all sets of lessons as interesting and of value. Some teachers appeared to prefer the contrast materials, with their inclusion of written tests; other teachers expressed a preference for the social perceptual lessons with their lesser demand for reading and writing. Their summary evaluations may be found on pages 5.17 and 5.18 of Chapter V.

In normal educational use, the restrictions on time for presentation and the restrictions on teacher-planned variations in use of the social perceptual lessons could be removed; lessons could be integrated into the year's full program and supplemented where desirable by more traditional academic tasks. Under normal circumstances in which lessons would be supplemented and repeated where necessary to achieve pupil mastery it seems reasonable to expect even greater and more consistent pupil gains in social cue comprehension.

V. TRIALS OF REMEDIAL MATERIALS AND RESULTS

A central project purpose was the development of training or educational procedures to improve the retardate's comprehension of the social situation and to extend the adequacy of his social behavior. Much project activity was involved with the experimental trial and subsequent revision of the training curriculum. In general format the training curriculum was a series of lessons that had been fully written out. Lessons were illustrated by a kit of slides and other materials so that minimal preplanning would be expected of the experimental teacher. The lessons were intended to illustrate for the pupil adult settings believed important to his independent functioning. Each such setting was discussed in relation to expected patterns of behavior. Each lesson, more evident in some than in others, attempted to clarify cues by which one could identify the situation and certain of its expected sequences of behavior. This material and the procedures for its development were described in the preceding chapter.

During the four years of project activities the training curriculum went through a number of revisions. An initial series of lessons (Curriculum I) received trial in two junior high special education classrooms; then was modified, based in part upon teacher evaluation, into Curriculum II, which in its turn was given trial in five institutional classrooms to gain information for its revision into Curriculum III.[32] The trial of Curriculum III in four junior high special education classrooms was the last trial of the program. A final revision of the lessons and accompanying materials, Curriculum IV, published with the title Social Perceptual Training for Community Living: Pre-Vocational Units for Retarded Youth, (Edmonson, Leach and Leland, 1967) represents a tangible product of this portion of the project emphasis.

To examine the effects of the successive revision of the social perceptual curriculum the trials involved a total of 11 separate classes of EMR adolescents or young adults. In each test situation the trial run of the curriculum was between six and ten weeks, with the additional time requirement of a week or more before and a week after, in which to complete the pre- and postevaluation of the subjects. The brief periods of experimental training required the teachers to "hurry through" the series of lessons with less than desirable regard for pupil assimilation and "carry over." The six week, eight week and ten week periods of training to be described were considered as minimal rather than optimal. The same less than optimal conditions obtained for the comparison classes.

In all but one of the trial situations, comparison treatment groups were used to control for factors of experimental involvement (Hawthorne effect) and unreliability of criterion measures. The following sections of this chapter describe each of these trials and their results. The overall assessment of the effectiveness of the experimental training to improve the retardate's social comprehension and the adequacy of his behavior is contained in the following chapter.

---

[32] An independent additional trial of Curriculum II was conducted in three public school junior high school classes in Nashville, Tennessee, as a dissertation project by Gary M. Clark (1967).

A. Trial 1

1. Procedures: The first controlled examination of the experimental curriculum involved five classes of special education pupils, four from the Shawnee Mission high school district in Johnson County, Kansas, and one from the adjacent Turner high school district in Wyandotte County, Missouri. The two districts had similar admission policies, including restrictions to pupils in the IQ range of 50 to 80. Each of the schools was several miles distant from the others and the experimental and placebo teachers, throughout, remained unaware of the participation of other teachers.[33] Near the end of the experimental period the control teacher had a brief description from an experimental teacher of the social perceptual materials he was using.

One control class and two experimental curriculum classes and two "placebo" treatment classes were planned. The single Wyandotte County class was selected as the control class, and the experimental and placebo classes randomly identified from the remaining four. The total school enrollment in the Turner School was approximately 600; in the Shawnee Mission schools the enrollments ranged from 956 to 1224. In size the Shawnee Mission schools resembled "University City School," a relatively large school described by Barker and Gump (1964, p 42 ff and Ch. 10) in their discussion of the inverse relationship they found to exist between social opportunities for pupils and the size of school and community. Both school districts, suburban, are part of the Kansas City metropolitan area. Residents are served by large shopping centers that must be reached by automobile. Public transportation is almost nonexistent. There are few sidewalks in these areas. Shawnee Mission families have higher incomes and tend to live in larger homes, in more compact neighborhoods, than is the case in the Turner district, which is rough and hilly and more sparsely settled. Average income is less in the Turner district.

The parents of Shawnee Mission pupils, more frequently than is the case in the Turner District, provide opportunities for their families to eat out in public, visit recreation settings, and travel during vacations. Shawnee Mission family acquaintanceships are probably larger than is the case in the Turner District. There are more working mothers in Turner families and there is less family time devoted to activities with their children.

The subjects were 54 mildly and moderately retarded adolescents whose age ranged from 13 to 16 years with a mean CA of 15 years. IQ's for these subjects ranged from 49 to 80 with a mean IQ of 66. Thirty-two subjects were boys; 22 were girls. The classes were closely similar as to IQ's and age as is indicated in Table 5.1.

---

[33] This was verified in discussion at a posttrial meeting attended by the five teachers and their coordinators.

TABLE 5.1

Means and Ranges of IQ's and Ages for Trial 1 Junior High School Retardate Experimental, Placebo, and Control (No Treatment) Classes, Singly and Combined

Group	N	IQ Mean	IQ Range	CA Mean	CA Range
Experimental Class A	11	69	51-80	180	151-199
Experimental Class B	8	64	55-71	181	171-199
Combined Experimental	19	67	51-80	181	151-199
Placebo Class C	15	67[a]	49-77	178	159-194
Placebo Class D	8	63	57-72	172	156-192
Combined Placebo	23	66[b]	49-77	176	156-194
Control Class E	12	70	60-79	178	147-196
Total	54	66	49-80	178	147-199

[a] N=14

[b] N=22

Two pre- and postexperiment measures were obtained from the 54 pupils involved in Trial 1, the Test of Social Inference (TSI) and the Teacher Behavior Rating Scales (T-BR).[34] The TSI was administered in individual sessions during the ten days prior to the experimental program, and during the week after its completion. Examiners did not know the group designations. Test protocols were coded so that groups could not be identified by test scores. A single TSI (Inf) score was computed for each subject for each of his two testings.

The observational and rating period for the weekly teacher behavior ratings (T-BR) was initiated three weeks before introducing the experimental, placebo, or control treatments to stabilize teacher impressions of pupils prior to the experimental period. The teachers continued to make weekly ratings throughout the treatment period and for two weeks after it ended. An intended third post-treatment behavior rating was precluded by the closing of the school year.

Lessons and visual materials for the final weeks of lessons were still being prepared as the trial began. The experimental teachers received the materials (Curriculum I) for the initial five weeks a week before their starting the program. The materials for the remaining three weeks were delivered as completed during the course of that month. The program consisted of lessons illustrated by 253 slides, tape recorder tapes and seatwork materials. It was practiced daily in sessions of from three-quarters of an hour to an hour's duration for an eight week period. Two visits were made by a project staff member to observe the classroom proceedings, once early and once late in the experimental period.

---

[34] All test instruments used in this and the succeeding trials reported in this current chapter are described in previous Chapters II or III.

The placebo teachers were provided with a selection of filmstrips and movies, with a time schedule for use during the eight week period. Both the placebo and experimental class teachers were told that they were participating in a project to evaluate audio-visual materials for later use with institutional pupils. Two observational visits to the placebo classes were made by the project staff.

The teacher of the control class was told her pupils would be given an experimental test. She was asked to evaluate pupils weekly on the T-BR behavior rating scales. She was given no additional information as to her role and until close to the end of the experimental period had no knowledge of the participation of other teachers. Toward the end of the treatment period in accidental meeting with an experimental teacher, she was given a brief description of the social perceptual material. Two observational visits were made to the control class by the project staff.

Material Evaluation forms were given to the experimental and placebo teachers with instructions to make ratings of their materials. On the form provided, the experimental teachers were to rate each lesson of Curriculum I with respect to pupil interest, language level, and content suitability. This scale permitted ratings "low," "below average," "average," "above average" and "high." On a similar form, the placebo teachers were to rate their filmstrips and movies with respect to the same criteria. No experimental activities were introduced in the control class and there were no interruptions to its program during the eight week period between the pre- and posttrial administration of the TSI and T-BR scales.

2. Results: (a) Test of Social Inference: TSI gain scores were computed by subtracting each pupil's pre-treatment TSI score from the score he earned after treatment (a nine week interval).[35] The mean TSI pretrial and gain scores for the participating experimental, placebo, and control classes and for the five classes combined are presented in Table 5.2. The two experimental classes and the two placebo classes were sufficiently alike (nonsignificant mean differences within the experimental and placebo groupings of both the TSI pre-trial and gain scores)[36] to permit combining these smaller classes

---

[35] Rosters of individual TSI and T-BR pretest and gain scores for all subjects in the several trials reported in this chapter are available at cost by writing to the senior author.

[36] The .05 level of confidence is used throughout this report.

into a single experimental and a single placebo group. The TSI pretrial and gain scores for these combined groups are also presented in Table 5.2.

TABLE 5.2

Means and Standard Deviations for TSI Pretrial and Gain Scores for Trial 1 Experimental, Placebo, and Control (No Treatment) Classes Singly and Combined

Group	N	Initial TSI Mean	Initial TSI SD	TSI Gain Mean	TSI Gain SD
Experimental Class A	11	55.4	22.1	12.2	5.6
Experimental Class B	8	50.2	23.7	15.0	17.2
Combined Experimental	19	53.3	23.0	13.4	11.6
Placebo Class C	15	57.7	17.4	7.7	8.1
Placebo Class D	8	44.5	18.7	6.2	6.5
Combined Placebo	23	53.1	18.9	7.2	7.4
Control Class E	12	55.2	15.4	4.2	6.3

As may be noted from the Table 5.2 data, the pretrial TSI means for the experimental, placebo, and control groups were considerably alike; means of 53.3, 53.1 and 55.2. The statistical equivalence of these three means was supported in terms of a non-significant *F* of .66. The mean TSI gains for the three groups, however, differed considerably, the mean gain of 13.4 TSI score points for the experimental group being nearly double that of the placebo group (mean gain of 7.2 TSI score points) and more than three times that of the control group (mean gain of 4.2 TSI score points); these yielding a significant between-group F of 4.48. Individual comparisons involving pairs of group mean gains yielded significant *t*'s of 2.24 and 2.80, and a nonsignificant *t* of .94 for the comparisons of experimental with placebo, experimental with control, and placebo with control, respectively. These between-group differences are perhaps even more revealing when the gains of individual subjects are examined. Only five of the 23 placebo pupils and only one of the 12 control pupils had gains as large as that of the average experimental pupil. Only three of the experimental pupils had gains less than that of the average control pupils, three additional experimental pupils having gains less than that of the average placebo pupil. Eleven of the 23 placebo pupils gained less than did the average control pupil.

In terms of TSI gains, most improvement in social cue interpretation was associated with the social perceptual curriculum. The second most improvement was associated with the placebo treatment of enriching the usual special education lessons with visual materials. The least (negligible) improvement was associated with the control (no experiment--nothing added) curriculum. This hierarchy of gain was viewed as supporting the project's general remedial approach to the retardate's social comprehension deficit.

(b) Teacher Behavior Ratings: As described in Chapter III teacher ratings of adequacy of their pupils' social behaviors were made using five social behavior scales permitting ratings from a low of one to a high of five. The five social behavior scales were social range, social relationship, peer acceptance, social invisibility and attentiveness. Three pretrial weekly ratings were averaged to provide an average pretreatment rating for each pupil on each of the five social behavior scales. Similarly, the posttreatment ratings that were assigned pupils in the last three rating weeks were combined to provide the average posttreatment ratings for each pupil for each of the five scales. Each pupil's change in rating (gain) on a given scale was the difference between his post- and pretreatment average rating on that given scale. The average pretreatment and gain Teacher Behavior Ratings for the five scales for the participating experimental and for the two placebo classes combined are presented in Table 5.3.

TABLE 5.3

Pretreatment Behavior Rating Means and Mean Gains on Five Rating Scales for Trial 1 Experimental, Placebo, and Control Classes, Separately and Combined

	Social Range		Relationship		Social Accept.		Social Invisibility		Attentiveness	
	Pre-Rat'g	Gain	Pre-Rat'g	Gain	Pre-Rat'g	Gain	Pre-Rat'g	Gain	Pre-Rat'g	Gain
Exper.Class A	2.5	+.1	2.2	+.5	2.4	+.5	2.4	+.8	2.1	+.8
Exper.Class B	2.1	0	2.1	+.1	2.2	-.1	2.2	+.4	1.8	+.1
Combined	2.3	+.1	2.2	+.4	2.3	+.3	2.3	+.7	2.0	+.5
Placebo Class C	2.9*	-.4	3.1*	-.9	3.2*	-.6	3.0*	-.4	2.7*	-.3
Placebo Class D	2.3*	+.3	2.4*	-.2	2.4*	-.2	2.2*	+.1	2.4*	-.2
Combined	2.6	-.2	2.9	-.6	3.0	-.5	2.7	-.1	2.6	-.2
Control Class E	2.5	+.8	2.7	+.2	2.7	+.2	2.9	+.1	2.4	+.6

* Placebo Classes C and D prerating means differ significantly at $< .05$ l.c.

The combinations of the two experimental and of the two placebo classes into single groupings of experimental and placebo subjects and the comparison of mean T-BR scores between these groups and the control class, using the analysis of variance design utilized for the TSI data analysis, was not followed for the rating data for two reasons: First, the pretrial ratings of the first placebo teacher were clearly (and significantly) higher than those of the second placebo class on all of the five rating scales, which prohibited meaningful combination of these two placebo classes. In addition, computation of product moment correlations for each of the classes to relate pupils' pre-ratings with their change or gain on the five behavior rating scales, revealed negative correlations in almost every instance. For the experimental classes the median correlation between preratings and gains for the five rating scales was -.17; for the placebo classes, the median was -.30; for the control class the median was -.75. Although on the one hand the sizable (negative) correlation between pretreatment ratings and gains suggested a covariance analysis to control for pretreatment difference, the variability of these correlations precluded the use of a single regression equation for adjusting the gain scores. In response to these conditions a more gross, nonparametric analysis was made of the teacher rating data. In order that the incidence of behavior rating gains be seen in relation to the magnitude of the initial rating, the median pretreatment rating of the total (54 pupil) sample was computed and pupils were dichotomized as either median and above (initially high) or below median (initially low) on each scale.

Fisher's exact probability test (Siegel, 1956, 96ff) was used as a test of significance of differences between treatment conditions with respect to number of pupils receiving higher posttreatment ratings, separately for both the initially high rated and the initially low rated pupils, for each of the five rating scales. In effect, a total of ten probability values were computed for differences between experimental and control, and ten additional for differences between placebo and control groups. None of the probabilities computed for differences between the experimental, placebo, and control groups in frequencies of higher post ratings were significant at the .05 level for the initially low rated pupils. For the initially high rated pupils, differences in frequency of gain between the experimental and placebo groups on the *social relationship*, *peer acceptance*, *social invisibility* and *attentiveness* scales all yielded significant p values. The difference in gains between the experimental and the control groups on the *social invisibility* scale also attained significance. One difference between the placebo and the control groups attained the .05 level--the greater gain by the control pupils with respect to *attentiveness*.

Summarizing these several comparisons, it may be concluded that the initially high rated pupils in the experimental group more frequently received higher posttreatment ratings than did initially high rated pupils in either the placebo or control groups. Differences between the experimental and control groups were generally smaller than was true for the experimental and placebo groups. For initially low rated pupils the data is different. No nonchance distinctions in frequency of increased teacher ratings following the treatment period were obtained between the three treatments for pupils initially rated low by their respective teachers.[37]

(c) Teacher Evaluation of Materials: The daily ratings by teachers of experimental and placebo materials with respect to Pupil Interest, Level of

---

[37] The idiosyncracy of the teacher-rater, as presented in Chapter III, was a later finding. On the basis of that more detailed analysis of ratings, the conclusion of difference favoring one treatment over another appears controvertible on the basis of T-BR data.

Language, and Suitability of Content, using the descriptive adjectives "low," "below average," "average," "above average," and "high," were converted to the numeric scale of 1 through 5 in order that the average of their ratings of each week's lesson might be computed. The values 3.0 - 3.9 represent assessments of "average." The average of the weekly ratings by the two experimental teachers and the two placebo teachers are listed below in Table 5.4. As can be noted, eleven of the 21 ratings of experimental materials were above "average." Eight of the ratings were average, and two were below average. These lower ratings were on Pupil Interest for the material prepared for the fifth and sixth week topics. Except for these two ratings, the teachers were positive in their evaluation of experimental curriculum with respect to Pupil Interest, Level of Language, and Suitability of Content. As can be noted, the placebo teachers also regarded their supplemental materials as appropriate to their own on-going curricula. Eight of the 24 ratings were above average. Thirteen of the ratings were average. Three were below average. Two of these lower ratings concerned the level of language in filmstrips used in the fourth and sixth weeks, a problem frequently encountered by special class teachers.

TABLE 5.4

Average Weekly Ratings[a] by Two Teachers Using Experimental Lessons and Two Teachers Using Placebo Audio-Visual Materials in Junior High School Special Education Classes

Week	Weekly Topics	Interest Holding	Language Level	Suitability Content
	Experimental - Social Perceptual Curriculum I			
1	Introduction to Signals	4.0	3.6	4.1
2	Places & Setting Tell Us What To Do	3.9	4.2	3.9
3	Field Trip	---	---	---
4	Signal Properties of Clothing	4.1	4.1	3.9
5	Where We Spend Our Time	2.6	3.5	3.4
6	People Having a Good Time	2.6	3.6	3.1
7	Best Social Behavior	4.4	4.0	4.0
8	Review	5.0	5.0	5.0
	Mean	3.8	4.0	3.9
	Placebo - Filmstrips and Films			
1	Personal Appearance Materials	3.8	3.5	4.6
2	Exercise, Health & Posture	3.0	3.8	3.9
3	Health & Safety	4.5	4.0	4.5
4	Baby Sitting and Auto Driving	3.0	2.7	2.8
5	USA and Transportation	3.7	3.4	3.8
6	USA "	4.2	2.8	3.0
7	USA "	4.5	4.0	4.1
8	Telephone and other Communication	3.0	3.5	3.5
	Mean	3.7	3.5	3.8

---

[a] Rating scale 1-5 (5=high)

The somewhat lower ratings for the fifth and sixth weeks' material in the social perceptual program were held to indicate the need for revision or elimination of those lessons in preparation of Curriculum II, the next edition of the experimental lessons.

B. Trial 2

1. Procedures: Subsequent to Trial 1, the curriculum was extensively revised. This revision involved a very general rewriting to reduce the level of language and add repetition.[38] Several lessons less well rated by Trial 1 teachers were replaced with new material. Supplemental lessons were added by which to illustrate several vocational settings.

The second experimental testing of the social perceptual training curriculum involved a tryout of this revision at the Parsons State Hospital and Training Center in Parsons, Kansas. The institution houses approximately 650 retarded nonadults between the ages of 6 and 21 and provides an extensive pattern of services and remedial programs for its residents with an emphasis on habilitation and preparation for community living. Approximately 60 patients are discharged to the community each year, principally from among those in the higher adaptive behavior groupings.[39]

Four classes of pupils were planned for Trial 2; two classes of boys and two classes of girls; one class of each sex group to receive the revised curriculum and the second class an alternative curriculum. To more adequately control for interclass differences and, especially, the possible differences between teachers in terms of their particular preference for either curriculum, the trial was replicated in a second session. In the second session, teachers and classes who first used the experimental curriculum then received the contrast curriculum, while teachers and classes who first used the contrast curriculum, then received the experimental curriculum.

The subjects were chosen from a listing of all older patients in the upper adaptive behavior levels (I and II) who were currently in the prevocational education program at the hospital. Prospective subjects were roughly ordered within sex groupings according to age and IQ and alternately assigned into one of two groups of nine subjects each. The groups of nine subjects were then randomly assigned to a same sex teacher, and then the teacher and his class were randomly assigned their curriculum treatment. The 36 subjects ranged in age from 14 to 20 years with an average age of 16 years, eight months. The IQ range for this sample was from 50 to 86 with a mean IQ of 70. The mean IQ's and ages of the four classes formed for the experiment were quite similar as is indicated in Table 5.5. The subjects ranged in grade level from 1.4 to 7.0 with 31 of the 36 subjects testing at achievement level above grade 2.6.

---

[38] See Chapter IV for a detailed description of this revision.

[39] All patients at Parsons are evaluated according to an Adaptive Behavior classification emphasizing the behavioral components of self help and personal and social responsibility (Leland, 1964)

TABLE 5.5

Mean IQ's and Ages for the Trial 2 Experimental and Contrast Classes, Singly and Combined

Class	Sex	N	Mean IQ[a]	Range	Mean CA	Range
Experimental Class F	F	9	72	61-84	198	179-214
" " G	M	9	70	59-83	201	171-240
Combined Experimental		18	71	59-84	200	171-240
Contrast Class H	F	9	74[b]	58-87	198	176-217
" " I	M	9	67	54-79	203	171-245
Combined Contrast		18	70[c]	54-87	200	171-245

[a] WISC or WAIS

[b] N=8 IQ datum is omitted for one pupil who tested below scale on two WISC subtests, but whose IQ was estimated from Draw-a-Man Test. Her inclusion would have reduced the mean to 72.

[c] N=17

The intended experimental design for assessing subject changes as an accompaniment of the social perceptual training required a comparison with subjects enrolled in an organized program of meetings of similar class time. The trial study at Parsons posed a special problem in this regard inasmuch as four new classes had been synthesized for the trial study; hence the two classes that would be used as comparison classes had no formal, on-going curriculum to be used as a control. In the new classes some pupils were new admissions to the institution; some had been called in from work placement; others had been reshuffled from existent special education classes. An alternative curriculum needed to be introduced; one which would be meaningful and supportive of the broader hospital prevocational program and at the same time conceal the experimenter's special interest in the social perceptual training curriculum.[40] The Occupational Educational Units (Fudell, 1963)[41] fulfilled these requirements by providing a competing curriculum with experimentally tested lessons which could be readily organized into a 10 week series of lessons. Several filmstrips and movies, as listed in Appendix F , were provided to illustrate the contrast curriculum, which, of itself, was nonvisual.

Both the contrast and the experimental curricula were introduced to the participating teachers as part of a study of the appropriateness of these two curricula. The experimental and contrast teachers were requested to use, criticize, and suggest improvements to their set of prevocational lessons. Teachers made evaluations daily. Teacher evaluation scales used in Trial 1 were expanded to include ratings of weekly lessons on Adequacy of Presentation, and estimation of Pupil Assimilation along with the original three criteria: Pupil Interest, Level of Language, and Suitability of Content.

---

[40] This concern that experimenter identification with a particular curriculum might bias teacher attitudes was especially important at Parsons where one of the project co-directors was an administrator at the hospital.

[41] These units are more fully described in Chapter IV. A list of lesson headings is included as Appendix E.

As with the Trial 1 procedures, both the Test of Social Inference (TSI) and the Teacher Behavior Rating Scales (T-BR) were used to provide pre- and posttrial data for assessing the effectiveness of the social perceptual training curriculum. The TSI was administered to all participating subjects in individual sessions the week prior to the experimental program, and again the week following its completion. This retest period was eleven weeks. The reversal of curriculum, by which the experimental curriculum teacher and classes received the contrast curriculum and the contrast curriculum teachers and classes received the experimental curriculum, commenced eight weeks later. The second[42] TSI testing served as a pretrial testing for this second round of curriculum trials. A third administration of the TSI followed completion of this second trial in early spring. A single examiner, who did not know pupil group assignments, administered the 36 pre- and the 36 posttests, and later, the additional retests. The examiner, a member of the Parsons Hospital staff, was trained in use of the TSI by the project staff. [43] All test protocols were coded so that groups could not be identified by test scorers.

Teacher ratings of their pupils on the five T-BR scales began with the inauguration of their experimental lessons and continued weekly for 10 weeks. The last training week of the first trial curriculum sessions coincided with the last week of the program. No immediate posttraining session ratings, as obtained for Trial 1 subjects, were possible since the classes were temporarily disbanded for the Christmas holidays at the completion of the 10 weeks training. The teacher ratings for the second round of curriculum trials commenced with that second session period and continued throughout for 10 weeks required for the training program. The final teacher ratings were collected the last week of the program.

In the week before the first training program began, the "contrast" teachers received their full 10 week packet of daily lessons and tests. At the same time the experimental teachers were furnished their designated lessons and the illustrative materials, for the initial five weeks. The remainder of the lessons and slides were delivered as completed by the project staff. During the 10 week training period two visits were made to each class by a member of the project staff, once in the initial week and once in the final week of the trials.

---

[42] This approximately two months' interim began with the forced interruptions of end of year (Christmas) home visits by pupils. The period was to permit staff respite from the demands of managing the four class curriculum and testing study.

[43] Appreciation is due Sandra Ward who administrated all of the pre- and posttests.

The "reverse" trial sessions were similarly conducted with minimal experimenter intervention: a single visit was made to each class. Not all the nine subjects were able to participate in the second trial. Eight subjects of the original 36 were unavailable for reasons of either hospital discharge, extended home visits, illness, or conflicting work assignments. Except for this difference of smaller class size, the procedures, lessons, and material were the same in both 10 week trial sessions.

2. Results: (a) Test of Social Inference. Two TSI gain scores were computed for each pupil, the first by subtracting his pretrial TSI score from his first posttrial TSI score, and the second gain score by subtracting his first posttrial TSI score from his second posttrial TSI score. The mean TSI pretrial and first and second gain scores for the participating experimental and contrast classes are presented in Table 5.6. Means for the combined boy and girl classes within each curriculum are also presented. The reduction from nine to six subjects per class for the "reversed" trials reflects the loss of as many as three subjects from two of the classes and a subsequent random removal of subjects' scores from the remaining classes to maintain the needed evenly sized groupings for the intended factorial analysis (see below). The second set of pretrial TSI means and the second posttrial gain scores are based only on these 24 retained subjects. The average total gain scores for each class reported in the rightmost columns of Table 5.6 are differences between the initial pretrial TSI score and the final, second posttrial testing for these same 24 subjects.

TABLE 5.6

Pretrial and Gain Scores from the Two Curricular Sessions
for the Trial 2, Means and Standard Deviations for the
TSI Experimental and Contrast Classes, Singly and Combined

	Experimental(N=9 S/class)				Contrast(N=6 S/class)				Total TSI	
	Pre TSI		TSI Gain		Pre TSI		TSI Gain		Gain	
Class	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
F-Girls	41.9	17.1	10.7	8.7	56.3	17.8	.7	8.8	10.3	7.4
G-Boys	48.1	16.2	20.7	8.4	69.8	15.1	1.5	9.7	24.8	11.4
Combined	45.0	17.0	15.7	9.6	63.1	17.8	1.1	9.3	17.8	12.1
	Contrast (N=9 S/class)				Experimental(N=6 S/class)					
H-Girls	42.7	18.1	13.2	7.9	50.2	14.5	6.2	10.3	20.3	15.3
I-Boys	43.2	7.9	11.7	8.6	54.6	13.0	7.3	5.0	16.7	4.5
Combined	42.9	14.0	12.4	8.3	52.3	13.9	6.8	8.1	18.4	11.4

As may be noted from the Table 5.6 data, the pretrial means for the four Trial 2 classes were generally close; means of 41.9, 48.1, 42.7, and 43.2. The statistical equivalence of these four means is in terms of a nonsignificant F of .27. In terms of their mean TSI gains following the first treatment session, all four classes showed sizeable score increases, the girls' and boys' contrast classes and the girls' experimental class gaining 13.2, 11.7, and 10.7 score points respectively. These gains are all quite similar. The experimental boys class, on the other hand, is alone with an average gain of almost 21 score points. These first sessions' gain scores were analyzed using a two-way analysis of variance design. Although the large gains by the experimental boys suggests an interaction effect, the considerable within-class variation in gains (differences between pupils in the same class of about 30 TSI score points in every class) resulted in a nonsignificant curriculum by sex F of 3.93. Similarly, both of the main factor effects, between curricula and between sexes, are interpretable as chance differences (F's of 1.22 and 2.10 respectively). The general conclusion from the 2x2 analysis is one of only chance differences between curricula and classes and between combinations of curriculum and class, differences no larger than would be expected more than one time in twenty to result simply from chance variations.

The second set of data from the "reversal" treatment (middle section of Table 5.6) based on classes of a third fewer subjects (eight not available and four whose scores were randomly discarded to maintain equal class sizes) reveals variable pretrial TSI means of 56.3, 69.8, 50.2, and 54.6 for the girls and boys, contrast and experimental classes, respectively, attributable to the differential treatment effects of the first 10 weeks' curriculum. Inspectionally, the TSI mean gains for this second treatment session clearly favor the experimental curriculum; negligible mean gains of .7 and 1.5 for the girls' and boys' contrast classes respectively, and mean gains of 6.2 and 7.3 for the girls' and boys' experimental classes, respectively. However, again the within class variability on individual gain scores vitiated other than random chance conclusions. The two-way analysis of variances for these second session TSI gain scores failed to yield significant F's for the between curricula comparison (F of 2.12). The F's for the between sex comparison and for interaction effects were both near zero, F's of .06 and .002 respectively. As with the analysis of the initial gain TSI scores, although mean differences favored the experimental curriculum, in view of the large individual variability within classes of pupils receiving similar curricula, these differences cannot be considered as reliably different from chance occurrence.

A more careful delineation of treatment and sex effects is possible by examining both sets of gain scores together following a Lindquist (1953) Type III mixed design. The results of this analysis are reported in Table 5.7, in which it may be noted that both the between curricula effect and the curriculum by order (of treatment assignment) interaction effects achieved significance at the .05 level. None of the other factors, order, sex, or the various interactions, were significant. The overall conclusion for the Trial 2 data reduces to one of significant, nonchance differences between mean gains in TSI scores favoring the experimental social cue training Curriculum II. The interaction effect suggests that the TSI gains are in part dependent on which curriculum, the experimental or the contrast, is received first. The data indicates that initial gains made following the contrast curriculum are continued with the experimental curriculum but not the reverse; that the initial gains in TSI score are greater starting with the experimental curriculum (at least for the boys) but that following the experimental curriculum with the contrast curriculum produces only minimal further gain.

TABLE 5.7

Analysis of Variance of TSI Gain Scores for the
Two Curricular Sessions for Trial 2

Source	df	SS	MS	F
Between	23	1687.00		
B (order)	1	19.42	19.42	.30
C (Sex)	1	108.83	108.83	1.65
BC	1	243.25	243.25	3.70
Error	20	1315.50	65.78	-
Within	24	3569.00		
A (curriculum)	1	359.42	359.42	4.36*
AB	1	1242.67	1242.67	15.06*
AC	1	262.26	262.26	3.18
ABC	1	54.82	54.82	.66
Error	20	1649.84	82.49	-
Total	47	5256.00		

* Significant at $\leq$.05 l.c.

(b) Teacher Behavior Ratings. As in the first trial of the social perceptual training curriculum, Teacher Behavior Rating (T-BR) scales were used to provide evaluation of pupil social adequacy. As noted earlier the four teachers rated their pupils at weekly intervals during both the experimental and contrast curriculum periods using the five scales from the T-BR; namely, social range, social relationship, peer acceptance, social invisibility and attentiveness.

The ratings on the first three of these scales were found so highly intercorrelated that rating on one was highly predictive of the rating on the others. As a simplification, therefore, ratings provided by the three scales were averaged into a single measure of combined social adequacy.[44] Following the procedures used in Trial 1, the pretrial ratings were computed as the average of the last two weekly ratings made during the given session. The pupil's change in rating (gain) on a given scale over the first session was the difference between his post- and pretrial average ratings for that treatment period. Similarly the pupil's rating gain over the second session was the difference between his post- and pretrial average ratings for that second treatment period. Table 5.8 presents the average pretrial ratings, average gains over the first session, average rating prior to the second session, and average gains over that period for the combined social adequacy, social invisibility, and attentiveness scales for each of the four classes involved in Trial 2. In the upper portion of the table the data are listed for the two classes that initially used the experimental materials and secondly used the contrast lessons. The lower portion of the table lists data for the two classes that initially used the contrast materials and secondly the experimental lessons.

---

[44] The correlations of ratings on the three contributing scales with the composite social adequacy scale rating were examined using the Parsons and Hissom (Trial 3) initial teacher ratings. These part-whole coefficients, each based on ratings of 70 pupils, were .82, .88 and .86 for the social range, social relationship and peer acceptance scales respectively. Correlations between the social invisibility ratings and the social adequacy composite rating and between the attentiveness ratings and the composite rating for these same pupils were .42 and .48 respectively.

Table 5.8 about here

As was true for the rating data reported for the five Trial 1 classes, teachers differed in their use of the rating scales. In addition, substantial negative correlations were obtained between pretrial rating and gain; that is, pupils initially rated low, gained the most and pupils initially rated high, gained the least. The nonparametric analysis procedures used for the Trial 1 rating data were therefore repeated here. These procedures consisted of combining all subjects into a single subject pool and identifying those subjects with pretrial ratings above the group median as "initially high" and those below the group median as "initially low" subjects. Fishers exact probability test (Siegel, 1956) was used as a test of significance of difference between classes with respect to the number of pupils receiving higher posttreatment ratings separately for the initially high rated pupils and the initially low rated pupils with respect to each of the three rating scales. The entire procedure was repeated in examining gains following the second curriculum session; the 24 second session subjects were dichotomized into high and low groups according to their average rating received at the end of the first session.

As might be expected from inspection of the generally small mean changes reported in Table 5.8, few significant interclass differences were found. With respect to the pupils whose initial ratings were low, none of the probabilities computed for differences between the experimental and contrast pupils on either curriculum session were significant. For the initially high rated pupils, the frequency of gain on the combined social adequacy scale was significantly greater for combined boy and girl pupils using the contrast materials for the first session. This is in contradiction to the expectation of greater improvement in social behaviors following the experimental curriculum.

In the second session, the only significant difference among rating gains was one between sexes. Independent of curriculum assignment, the initially high boys gained more frequently than initially high girls on the attentiveness scales. On none of the three rating scales were differences between the frequencies of gains in ratings made of them during the second curriculum

TABLE 5.8

Pretreatment Behavior Rating Means and Mean Gains on Rating Scales for Trial 2 Classes Who Used Experimental and Contrast Materials in Opposite Sequences

	Social Adequacy				Social Invisibility				Attentiveness			
	Session 1		Session 2		Session 1		Session 2		Session 1		Session 2	
Class	Pre-Rat'g	Exp. Chg.	Pre-Rat'g	Con. Chg.	Pre-Rat'g	Exp.[a] Chg.	Pre-Rat'g	Con.[b] Chg.	Pre-Rat'g	Exp. Chg.	Pre-Rat'g	Con. Chg.
F-Girls	2.5	-.2	2.4	+.1	2.2	-.4	2.2	-.4	2.2	0	2.4	-.1
G-Boys	2.2	-.2	2.2	+.2	1.9	+.2	2.2	+.2	2.6	+.1	2.4	+.1
Combined F + G	2.5	-.2	2.3	+.2	2.0	-.2	2.2	-.2	2.3	0	2.4	0
		Con. Gain		Exp. Gain		Con. Gain		Exp. Gain		Con. Gain		Exp. Gain
H-Girls	2.4	+.1	2.6	-.2	2.8	0	3.0	-.7	2.9	-.5	3.1	-.2
I-Boys	1.7	+.5	1.9	-.2	2.0	+.2	2.0	+.1	1.8	+.1	1.5	0
Combined H + I	2.0	+.3	2.2	-.2	2.4	+.1	2.5	-.4	2.4	-.2	2.3	-.1

Note: Experimental curriculum used in Session 2 for Classes F and G and in Session 2 for classes H and I. Contrast curriculum used in Session 2 for Classes F and G and in Session 1 by classes H and I. Nine S's each class for Session 1. For Session 2, classes F and G had 8 S's; classes H and I had 6 S's.

[a] Experimental

[b] Contrast

session different for the experimental and contrast groups. The general picture of interclass comparisons of changes in pupil social behavior as measured by pre- and postcurriculum teacher ratings is simply one of no differences.[45]

(c) Teacher Evaluation of Materials. During both curriculum sessions the teachers made daily evaluations of the experimental and contrast lessons with respect to Pupil Interest, Level of Language, Suitability of Content, Adequacy of Presentation, and Pupil Assimilation. The descriptive adjectives on the evaluation form of "low," "below average," "average," "above average," and "high" were converted to a numeric scale of 1 through 5. Particularly evident in Trial 2 were differences between teachers in their daily evaluations of the lessons. This was true for both experimental and contrast lessons but was especially apparent in session 1 with respect to the contrast materials. This difference in teacher attitude toward the materials may have been a variable in relation to pupil TSI gains. As an example, the teacher of the first session boys' class using the contrast lessons gave 73 "below average" ratings of his total of 175 ratings (42 percent). Only five of his ratings were "above average" (3 percent). On the other hand, the teacher of the first session girls' class evaluated none of the contrast lessons as "below average." She gave 101 "above average" ratings of her total of 125[46] ratings (81 percent). The teachers using the experimental curriculum in the first session were in closer agreement in their evaluations of their material. The girls' teacher assigned 158 of her 219 ratings (72 percent) to the "above average" and 11 (5 percent) to the "below average" categories. The boys' teacher assigned 154 of his 220 ratings (70 percent) to the "above average" and 2 (1 percent) to the "below average" categories.

In the second session the boys' class teacher who had disliked the contrast materials was more favorable in his evaluation of the experimental materials. Sixty-two percent of his ratings were "above average"; four percent were "below average." The girls' class teacher who had liked the contrast materials was much less favorable in her evaluation of the experimental materials. Thirty-three percent of her ratings were "above average", fourteen percent were "below average." The two teachers who had closely agreed in their evaluation of the experimental lessons diverged in their evaluation of the contrast lessons. The teacher of the boys' (now) contrast class rated the materials mostly "average"; forty-one percent of his ratings were "above average"; one percent were "below average." The teacher of the girls' (now) contrast class rated the materials mostly "above average." Ninety-seven percent of her ratings were assigned to that category. None were "below average."

One of the male teachers in a tape recorded post-trial conference expressed his preference for the experimental lessons in these words:

---

[45] See Footnote on page 5.7

[46] Misunderstanding concerning the evaluation schedule resulted in no evaluation of lessons for weeks 7, 8 or 10 by one teacher and no evaluation for weeks 5, 7 and 10 by the second teacher.

> I don't think there is any question about the value of the material... I think they were good, both the first one and the last one. I think my problem with the first one /contrast lesson/ was...too many of my boys can't read...I think the second one /social perceptual curriculum/ seemed to generate more interest because they didn't need to read and there were more visual aids.

One of the women teachers expressed her preference for the contrast materials:

> Personally, I liked the first one /contrast/ better and my girls seemed to like it better. These girls are pretty high level girls and the second one /experimental/ seemed pretty infantile to them...I think the level of the material was too low...They were surprised by the lack of reading and writing...

In summary, the male teachers of the institutional boys' classes preferred the experimental lessons. The women teachers of the girls' classes preferred the contrast lessons. These differences in teacher attitude toward the two sets of lessons may have had an effect upon pupil performance which cannot well be isolated in the statistical presentation. It was evident to the staff observer who visited each class on a total of three occasions (spread over the two sessions) that the weekly written tests,[47] in connection with the contrast lessons, were viewed by the male teachers as inappropriately difficult for the boys. On the other hand, the lack of written assignments in connection with the experimental materials was viewed as somewhat inappropriate by teachers of the girls' classes. These differences in evaluations are viewed as endorsing the value of both sets of materials, while also reflecting reasonable differences in opinion as to the suitability of the two sets of lessons for different groups (or for use by different teachers).

C. Trial 3

1. Procedures: The third experimental testing of the social perceptual training program involved a further tryout at the Hissom Memorial Center at Sand Springs, Oklahoma of the same revised curriculum introduced at Parsons (see B, above). Hissom Center, then in its third year of operation as a state hospital for the retarded, was serving approximately 600 boys and girls between the ages of 4 and 26 years whose IQ's ranged from 83 downward. In keeping with the broad hospital goal of developing independent living skills, extensive in-hospital and off-campus activities were provided at the trainable and educable levels. A part of these activities were the six special education classes initiated at the hospital the preceding fall in which 158 pupils were enrolled.

The 34 pupils selected for a further trial testing of the social perceptual training program were all enrolled in two of these classes. The classes represented two ranges of educational level, a sixth through eighth grade class and a fourth through sixth grade class, and in effect constituted the total available population of prevocational level Hissom residents.

---

[47] The tests were prepared for pupils having second grade reading ability. The average CAT grade level for the boys was 4-7. One boy for whom CAT data was lacking may have been at grade level below the second grade. Three boys were at grade levels between 2-0 and 2-6. The average CAT grade level for the girls was 3-11. Two girls were at grade levels below 2-0; one additional girl was at grade level below 2-6.

The pupils from these two classes were regrouped into same-sex pairs roughly equated with respect to IQ and age. Grade level was not considered. Two extreme subjects unmatchable in terms of age were dropped from the sample. One pupil from each pair was then randomly reassigned to one of two classes of ten girls and seven boys. The 34 retained pupils ranged in age from nearly 12 years to 19 years, with a mean of 15 years, nine months. Their IQ's ranged from 43 to 83 with a mean of 62. Fourteen pupils were boys; 20 were girls. The pupils averaged 11 months younger than the Parsons' subjects in Trials 2 and 3. Their average IQ was 8 points lower than the Parsons' average. The ranges and means of IQ's and ages for the boys and girls in each of the classes and for the total classes are presented in Table 5.9.

TABLE 5.9

Means and Ranges of IQ's and Ages for Girls and Boys in the Trial 3 Experimental (J) and Contrast (K) Classes

		IQ		CA	
	N	Mean	Range	Mean	Range
Girls	10	61	48-77	182	148-216
Boys	7	62	43-77	199	175-232
Class J	17	61	43-77	189	148-232
Girls	10	64	48-83	191	141-216
Boys	7	62	46-72	189	168-219
Class K	17	63	46-83	190	141-219
Total	34	62	43-83	189	141-232

The Hissom trial replicated the first part of the Parsons trial in that one class was randomly designated to receive the social perceptual training program and the second to receive the Occupational Educational Units as a contrast curriculum. Aside from this basic similarity, however, there were several differences between the two trials. For one thing the classes were larger (N=17) than the Parsons classes (N=9) and contained both boys and girls. Also, as already noted, the Hissom subjects were nearly a year younger than the Parsons subjects on the average and averaged lower with respect to IQ's. Because of management-material problems no added visual materials were provided for use with the contrast curriculum as had been the case at Parsons. A final difference was the use of a single teacher teaching both classes, the contrast class in the morning and the experimental class in the afternoon. Lessons for both classes were used daily for a period of 10 weeks. The teacher was instructed that she was to use, criticize, and suggest improvement to two sets of experimental materials. To facilitate her comments, she was provided with tape recorder tapes.

Pupils were administered the TSI in individual sessions the week prior to use of the curriculum. Most pupils were retested the week after completion of the lessons. A few pupils had their retests delayed by a long weekend holiday. No difference in average TSI score was obtained between those retested before and those retested after the weekend holiday. All tests were administered by nearby University students who had been trained in TSI administration by the project staff. Neither examiners nor scorers were aware of group designation.

As with the prior trials, the Teacher Behavior Rating forms (T-BR) were used to provide social adjustment ratings of all pupils. These ratings were made on each of five scales, *social range*, *social relationship*, *peer acceptance*, *social invisibility*, and *attentiveness*, by the classroom teacher once during the week prior the treatment period and once during the week after the final lesson.

In addition, an "outside" measure of pupil social competence was obtained on all pupils. The "outside" raters were six staff persons well acquainted with the 34 pupils: the chaplain, school secretary, librarian, nurse, physical education teacher, and director of cottage life. After participating in a recall discussion of retarded persons he had known, each rated the pupils using the Outside Behavior Rating (O-BR) form and procedures described in Chapter III above. These ratings, to be based on nonclassroom pupil behaviors, were made once before and once immediately following the treatment period.

2. Results: (a) Test of Social Inference: TSI gain scores were obtained by subtracting each pupil's pretreatment TSI score from the posttreatment score he earned eleven weeks later. The mean TSI score, the mean gain score, and standard deviations for both scores are presented for both classes in Table 5.10.

TABLE 5.10 about here

As may be noted from Table 5.10, the pretrial TSI means for the boys in the experimental and contrast classes differed by approximately seven TSI score points, and for the girls by less than two TSI score points. Examination of these differences in terms of *t* tests revealed that neither was significantly different from chance expectation; *t*'s of .89 for the boys and .12 for the girls.

TABLE 5.10

Means and Standard Deviations for the TSI Pretrial and Gain Scores for Girls and Boys in Trial 3 Experimental (J) and Contrast (K) Classes

Sample	N	Pretrial TSI Mean	Pretrial TSI SD	TSI Gain Mean	TSI Gain SD
Girls	10	38.4	27.1	11.2	6.5
Boys	7	34.0	11.0	17.4	14.1
Class J	17	36.6	22.8	13.8	10.8
Girls	10	37.1	19.7	10.6	9.8
Boys	7	27.3	14.8	2.3	5.2
Class K	17	33.1	19.1	7.2	9.2
Total	34	34.8	20.8	10.5	10.7

The TSI gain scores noted in Table 5.10 were examined in terms of a 2x2 analysis of variance design following the Rao interaction procedure (Rao, 1952) correcting for unequal cells. With neither the between treatment effects (F=3.77) nor the between sex effects (F=.12) significant at the .05 level, the significant treatment by sex interaction effect (F=4.39) supports a further examination of between treatment differences in gain scores in terms of the two sex groups separately. The significant t ratio of 2.87 obtained for the boys indicates a nonchance difference of greater TSI score gains made by retarded boys receiving the experimental curriculum. In terms of the TSI gains reported in Table 5.10 this difference is one of 15.1 TSI score points. This support for the experimental curriculum did not extend to the girls, however; the small mean difference in TSI gains of .6 and the t of .13 indicating minimal difference between the two curricula for the retarded girls.

(b) Teacher Behavior Ratings. As with the Parsons data analysis, the five teacher rating scales used in rating the experimental and contrast pupils were collapsed into three scales providing a social adequacy rating (an average of ratings made on the social range, social relationship, and peer acceptance scales), a social invisibility rating, and an attentiveness rating. The initial pretrial rating means for both the experimental and contrast class girls and boys, separately and combined, are presented in Table 5.11 together with mean gain scores computed as differences between the pretrial and posttrial ratings.

TABLE 5.11

Pretreatment Teacher Behavior Rating Means and Mean Changes on Three Rating Scales for Girls and Boys in Experimental and Contrast Classes

Sample	N	Social Adequacy Mean Pre-Rat'g	Social Adequacy Mean Change	Soc.Invisibility Mean Pre-Rat'g	Soc.Invisibility Mean Change	Attentiveness Mean Pre-Rat'g	Attentiveness Mean Change
Experimental							
Class J							
Female	10	1.7	+1.0	2.9	+.3	2.2	+1.3
Males	7	2.4	+ .8	3.1	0	2.1	+1.4
Combined	17	2.2	+ .9	3.0	+.2	2.2	+1.4
Contrast							
Class K							
Female	10	2.2	+1.2	3.6	+.3	3.0	+ .6
Males	7	1.6	+1.6	2.0	+.6	1.3	+1.6
Combined	17	2.0	+1.4	2.9	+.4	2.3	+1.0

The nonparametric procedures followed for the T-BR analysis for Trial 1 and 2 were again followed for the Hissom data.[48] Briefly, the procedure involved examining the differences in relative frequencies of experimental and contrast rating separately for pupils identified as initially high ratees, who received still higher posttrial ratings, and separately for those remaining pupils identified as initially low ratees who received higher ratings at the close of the program. Only one of these differences was significant in terms of Fisher's exact probability test; more of the initially low rated pupils in the contrast class gained with respect to teacher ratings of _social invisibility_ than did the initially low rated pupils in the experimental class. The fact that this difference favors the contrast curriculum and that none of the other five comparisons was significant suggests no advantage of the social perceptual training program over the Occupational Educational Units in effecting behavioral adjustment changes in the pupil as rated by his teacher.[49]

(c) Outside Behavior Ratings: The experimental and contrast class pupils were further compared in terms of their out-of-class behaviors as rated by the six nonclassroom raters using the same O-BR format. These comparisons are reported for the same three social behavior dimensions rated by the teacher, namely _social adequacy_, _social invisibility_, and _attentiveness_.

---

[48] See p. 5.7 for a more detailed description of this analysis.

[49] See footnote on page 5.7.

As noted earlier (Chapter III) substantial interrater agreement was generally obtained for these ratings, corrected[50] split half correlation coefficients of .87, .90, and .89, for the social adequacy composite scale, invisibility and attentiveness, respectively. Moderately low correlations, between the teacher ratings and the "outside" behavior composite rating (r's of from .33 to .57) supports this further examination of possible behavior change related to the experimental curriculum.

The O-BR gains were computed similarly to those for the T-BR; that is, as differences between the average outside behavior rating received by each pupil before and after the curriculum session. Pretrial rating means based on the six outside raters using the three O-BR scales and mean gains or changes in these ratings are reported in Table 5.12 for the experimental and contrast Hissom pupils in terms of both class means and in terms of means for sex subgroups within classes.

TABLE 5.12

Pretreatment Mean Ratings and Mean Changes in Ratings Made by Six Outside Raters of Social Behavior of Boys and Girls in Trial 3 Experimental and Contrast Classes

		Social Adequacy		Soc.Invisibility		Attentiveness	
	N	Mean Pre-Rating	Mean Change	Mean Pre-Rat'g	Mean Change	Mean Pre-Rat'g	Mean Change
Experimental							
Class J							
Girls	10	2.1	+.2	1.9	+.3	2.0	+.4
Boys	7	2.3	+.4	2.4	+.4	2.4	+.2
Combined	17	2.2	+.3	2.1	+.3	2.1	+.3
Contrast							
Class K							
Girls	10	2.4	+.4	2.6	+.4	2.7	+.3
Boys	7	2.2	+.2	2.2	+.2	2.2	+.2
Combined	17	2.3	+.3	2.4	+.3	2.5	+.2

The analysis of Table 5.12 pretrial and gain scores on the three O-BR rating scales was identical to that for the previously reported teacher rating scale data. Fisher's exact probability test was used to determine the significance of differences in frequencies of pupils from the experimental and contrast classes who received higher posttrial ratings. None of the between-class differences achieved significance at the .05 level. As with the T-BR data, argument for a greater positive effect of the social perceptual training program on the pupil's social competence than that of the contrast Occupational Educational Units found no support in this trial, as measured by the nonclassroom observers using the O-BR.

---

[50] For double test length using the Spearman Brown formula

In this trial the teacher tape-recorded her evaluations of the materials. She had specifically been asked to criticize the materials and make suggestions for their improvement. A number of her suggestions were responded to in the next revision of the social perceptual materials.[51]

D. Trial 4

1. Procedures: Following the Hissom trial a very extensive revision was made of the social perceptual training curriculum. A major objective of this revision was to reduce the apparent sex bias inadvertently built into Curriculum I and II lessons and material as indicated by higher TSI gains by boys during all foregoing trials, but especially apparent in Trials 2 and 3. A girl was introduced into the narrative of certain of the lessons as a second central figure. The vocational setting lessons in Curriculum II were replaced with new lessons affording a broader "social mobility" orientation, with emphasis on social interaction and leisure time activities. Although teachers and pupils had responded favorably to the vocational setting lessons, they were replaced by other lessons so as to supplement or complement the contrast curriculum rather than to duplicate it or overlap it.

The lessons were all completely rewritten using less repetitious, higher level, language than in the previous revision.[51] The limit was retained of one lesson a day for 10 weeks.

This second revision of the lessons (Curriculum III) was trial tested in four classes of junior high special education pupils, two classes in Springfield, Missouri and two located in the Shawnee Mission, Kansas City, Kansas, area. The two Springfield schools (L and M) were located in lower socioeconomic neighborhoods in a city of approximately 100,000 residents. They enrolled approximately 350 and 650 pupils, respectively. There were few Negro families represented in either of the two neighborhoods. The families of School L were described by the school principal as being longtime residents, as maintaining steady employment, and as providing reasonably good home environments as a rule. The families of School M were described as being more transient, as having a higher rate of unemployment, as requiring more social welfare assistance, and as providing somewhat less supportive home environment, as a rule. Pupils in both schools were characterized as being rather self-reliant and as getting around in the city by foot, by bicycle, or by bus.

The Shawnee Mission District was described under Trial 1 as a suburban area of metropolitan Kansas City. School N is located in a new above-middle socio-economic area. School O is located in a much older area that is more heterogeneous with respect to family occupations and income. The pupil enrollments in these schools were approximately 1025 in School N and approximately 980 pupils in School O. The schools were almost exclusively white.

---

[51] These revisions, based in part on user's (teachers) recommendations obtained from previous trials, are described in greater detail in Chapter IV.

The special education classes in Schools L and M were divided on the basis of chronological age as "7th grades" and "8th grades." In School L the intact "8th grade" special education class provided 14 subjects for trial of the materials. The teacher was an experienced teacher who was teaching a special education class for the first time. The pupils were 8 boys and 6 girls with ages ranging from 13 to 16 years with a mean of 15 years. Their IQ's ranged from 62 to 85 with a mean IQ of 72. Most of the pupils enrolled in the seventh and eighth special education classes in School M were boys. To increase the number of available girls, these two grades were combined into a special class meeting each day for the period during which the experimental curriculum was presented. The class was conducted by the 8th grade special education teacher assisted by the seventh grade teacher. Both teachers were experienced teachers. The 8th grade teacher had had long experience in special education. The pupils were 16 boys and 8 girls with ages ranging from 12 years five months to 15 years, with a mean of 14 years. Their IQ's ranged from 52 to 88 with a mean of 76.

The two Shawnee Mission classes provided a total of 23 subjects. The 12 pupils in School N were 5 boys and 7 girls with ages ranging from 12 years to 16 years with a mean age of 15 years. Their IQ's ranged from 56 to 82 with a mean IQ of 71. The 11 pupils in School O were 8 boys and 3 girls with ages ranging from 12 years to 15 years with a mean age of 14 years. Their IQ's ranged from 45 to 98,[52] with a mean IQ of 70. The mean IQ's and ages of the four classes of Trial 4 subjects are presented in Table 5.13. The teachers of classes N and O were completing their first year of teaching. Similar data for Class E, the no treatment class used as a control class in Trial 1, is also included in this table. Data from this latter class was reused for the current trial analysis to provide a group baseline for examining pre-post treatment changes in experimental Classes L, M, N, and O.

TABLE 5.13

Mean IQ's and Ages for the Trial 4 Experimental Classes, Singly and Combined, and for the Trial 1 Control Class

Class	N	IQ	Range	CA	Range
Experimental Class L	14	72	62-85	179	166-195
Experimental Class M	24	76	52-88	166	149-183
Experimental Class N	12	71	56-82	169	153-202
Experimental Class O	11	70	45-98	166	152-186
Combined Experimental	61	73	45-98	169	149-202
No Treatment Class E	12	70	60-79	178	147-196

---

[52] The "98 IQ" pupil was so listed on the special class roster. He was found to score on the TSI as a retarded subject (initial TSI score of 52). He was noted by the examiner as being very reticent.

The testing of the revised curriculum in the four selected special education classes ran into a serious timing problem. By the time the revision of the social perceptual training lessons had been completed, the school year was too near its end to allow the sequential trial of the complete 10 week series of 50 lessons. Only six usable weeks remained. To permit a trial use of all lessons, although at the cost of disrupting curriculum lesson sequence and the repetition of concepts, the revised curriculum was shortened to two partially overlapping "halves," one half consisting of lessons for weeks 1, 2, 4, 6, 8 and 10 ("even numbered") and the second half consisting of lessons for weeks 1, 3, 5, 7, 9 and 10 ("odd numbered"). One teacher in each school district was randomly selected to receive the "even numbered" lessons, the "odd numbered" lessons being assigned to the second teacher in that district. Teachers in Schools L and N used the "even numbered" lessons; teachers in Schools M and O used the odd numbered lessons. Altogether, pupils in each class were exposed to thirty lessons in daily meetings over a period of six weeks.

As in previous trials, both the TSI and the T-BR were used in assessing the effects of the social perceptual training curriculum. Pupils were administered the TSI in individual sessions the week before beginning use of the experimental lessons and were retested seven weeks later at the conclusion of the program. Tests of L and M pupils were administered by students of a local college who were trained in use of the TSI by the project staff. Tests of N and O pupils were administered by the project staff. In all schools the special education teachers rated their pupils on the five T-BR scales the week prior to use of the lessons and rerated them the week after its completion. As described under Trial 1, the control class pupils were simply tested and rated immediately before and immediately following an eight week period of their regular classroom activities, using both the TSI and the T-BR scales.

2. Results: (a) Test of Social Inference. TSI gain scores were computed for each pupil in Classes L, M, N, and O by subtracting his pre-test from his post-test score. The average TSI pretrial score, the average gain score, and standard deviations for both scores are presented in Table 5.14 for the four Trial 4 classes and for the no treatment control class. Because of experimenter interest in between-sex groups as well as between-class comparisons, the data are presented for girls and boys separately within classes, for each boys-plus-girls class, for the combined experimental girls as a group, for the experimental boys as group, and for the total combined four experimental classes.

Table 5.14 about here

As may be noted in Table 5.14, the pretrial TSI means vary somewhat among the experimental classes, class means ranging from 60.1, 55.1, 52.8, and 44.0. The two most extreme class means are in large part accountable for by singularly deviant pupils, one boy in Class N earning an unusually high TSI score of 104, a full standard deviation above that earned by most nonretarded adolescents (the mean TSI scores reported for 45 nonretarded junior high school pupils was 86 with a sigma of 15 TSI score points)[53] and one boy in Class O earning an equally unusually low TSI score of 8, approximately two sigmas below that of the obtained class mean for educable retardates. The individual extreme scores resulted in larger standard deviations for classes N and O.

TABLE 5.14

Means and Standard Deviations for the TSI Pretrial and Gain Scores for Girls and Boys in Four Trial 4 Experimental Classes and for a Control Class

Class	N	Initial TSI Mean	Initial TSI SD	TSI Gain Mean	TSI Gain SD
Experimental Class L Girls	6	52.1	8.6	11.8	7.3
Experimental Class L Boys	8	57.4	14.4	11.9	5.9
Experimental Class L Combined	14	55.1	11.7	11.9	6.0
Experimental Class M Girls	8	45.1	8.1	7.4	3.9
Experimental Class M Boys	16	56.7	14.2	12.2	8.3
Experimental Class M Combined	24	52.8	13.5	10.6	7.4
Experimental Class N Girls	7	59.0	14.8	1.8	5.0
Experimental Class N Boys	5	61.6	27.1	11.2	10.5
Experimental Class N Combined	12	60.1	18.8	5.8	8.4
Experimental Class O Girls	3	42.7	17.9	8.7	4.2
Experimental Class O Boys	8	44.5	22.8	5.6	6.1
Experimental Class O Combined	11	44.0	19.8	6.4	5.4
Experimental Girls Combined	24	50.6	12.7	7.0	6.2
Experimental Boys Combined	37	54.8	18.4	10.6	7.9
Experimental Classes Combined	61	53.1	16.4	9.2	7.4
No Treatment Class E Girls	6	61.3	12.9	1.2	4.3
No Treatment Class E Boys	6	49.0	13.7	7.3	4.8
No Treatment Class E Combined	12	55.2	15.4	4.2	6.3

In anticipation of three intended comparisons of the TSI gain scores, (1) a comparison of gains for the experimental and control classes, (2) a comparison of gains for the experimental class girls with those for the experimental class boys, and (3) a comparison of pupils in those classes receiving the odd numbered lessons, possible distorting differences between the three, and the even groupings of classes to be involved in these comparisons in terms of their initial and pretrial TSI scores needed to be examined. With respect to the first, experimental classes vs. control class, comparisons, it was proposed to combine the four experimental classes into a single group for comparison with the single control class. To examine the appropriateness of this proposed combination, the initial TSI scores for the four classes were compared in terms of a simple across-groups analysis of variance design. The obtained F of 2.01 was not significant at the .05 level, indicating that mean differences in TSI pretrial scores as large as those noted in Table 5.14 are within the range of chance variations and that combining of these classes for comparisons with the control (no treatment) class is statistically reasonable. It will be noted here that the mean TSI score of 53.1 for the combined four classes is also reasonably close to that for the control class pretrial TSI mean of 55.2.

---

[53] See Table 2.3 on page 2.10.

Differences in initial TSI scores for the experimental girls and boys were compared in terms of *t*-ratio. As may be noted from Table 5.14, the difference in the initial TSI means for the experimental girls and boys was 4.2 TSI score points in favor of the boys. The *t* ratio for this difference of .98 indicates only a chance sex difference in initial TSI score means for the experimental pupils.

Differences in initial TSI scores for those experimental classes receiving the even numbered lessons (Classes L and N) and those classes receiving the odd numbered lessons (Classes M and O) were similarly compared in terms of a *t* ratio. As may be computed from the class means in Table 5.14, the combined classes L and N initial TSI mean was 57.42, that for the combined classes M and O was 50.06. The mean difference between these two combined classes of 7.36 TSI score points yielded a *t* of 1.76, indicating only chance differences between the initial TSI score means for experimental pupils receiving the two differently abridged curricula.

The difference in TSI gain scores for the combined four class experimental group and for the no treatment control class was examined in terms of a *t* ratio. The mean gain for the experimental pupils after six weeks of social perceptual training sessions was 9.2 TSI score points; for the control class the gain was 4.2 TSI score points, over a slightly longer, eight week, non-experimental treatment period. The *t* ratio for the difference in gains between the experimental group and the no treatment class was 2.13, significant at the .05 level, indicated the gain in TSI performance associated with the six week experimental curriculum session was beyond that usually attributed to chance variation. The gain by experimental class pupils, although not as great nor as uniform as gains in previous trials, was achieved within a shorter span of training and in spite of the discontinuity introduced into the experimental program with the separation of the material into "odd numbered" and "even numbered" lessons.

The comparison of the performance of the girls and boys in terms of gain scores on the TSI retest after treatment was similarly made in terms of a *t* ratio. As reported in Table 5.14, the gain score for the 23 experimental girls was 7.0 TSI score points, for the 38 experimental boys 10.6 points. The difference of 3.6 points yielded a nonsignificant *t* of .60, indicating only chance differences between the boys and girls with respect to TSI performance subsequent to the revised social perceptual training curriculum.

The comparison of performance of special education pupils who received the two different portions of the revised curriculum similarly yielded a nonsignificant ratio (*t* of .12), indicating that the obtained difference between the combined L and N Class pupils and M and O Class pupils of .24 was well within limits of chance expectancy. This very similar gain by pupils receiving either segment of the training curriculum suggests a general homogeneity of the lessons with respect to performance as measured by the TSI.

(b) Teacher Behavior Ratings. The differences between each pupil's pre- and posttraining ratings on the T-BR scales were computed as his behavior rating change scores on these scales in Table 5.15. The mean preratings and the mean changes are presented for each of the four experimental classes separately and combined.

TABLE 5.15

Pretreatment Behavior Rating Means and Mean Change in Rating for Trial 4 Experimental and Control (No Treatment) Classes

		Social Adequacy		Social Invisibility		Attentive-ness	
	N	Pre-Rating	Change	Pre-Rating	Change	Pre-Rating	Change
Experimental Class L	14	2.4	+.3	2.3	+.6	1.6	+.6
" " M	24	1.5	+.5	1.8	+.6	1.6	+.2
" " N	12	2.5	-.3	2.8	+.3	2.7	-.6
" " O	11	1.9	+.1	2.0	-.3	1.5	+.2
Combined Experimental	61	1.9	+.2	2.1	+.4	1.8	+.1
Control Class E	12	2.6	+.4	2.9	+.1	2.4	+.6

As may be seen from these entries, the teachers apparently varied considerably with respect to their ratings of their pupils as higher or lower at the end of the training session. In the combined group there was little indication of higher ratings.

The detailed analysis of teacher ratings which is presented in Chapter III casts doubt upon the validity of changes in T-BR ratings. As is described in that chapter the extreme differences in use of the scales by various teachers, in combination with the strongly inverse correlation between initial ratings and changes in ratings obtained in all classes, leaves interpretation of this data in doubt. However, following the same nonparametric procedures used for the T-BR analysis in Trials 1 through 3, the T-BR "gains" were examined separately for pupils whose initial ratings were identified as high (at the group median or higher) and for those remaining pupils identified as initially low ratees (rated lower than group median). Difference in frequency of gain between the initially high rated experimental pupils and the initially high rated control pupils was examined in terms of Fisher's exact probability test. A similar comparison across groups was made of gains by the initially low rated pupils. None of the difference in gain frequencies exceeded chance levels.

E. Summary

It was hypothesized that social comprehension could be taught the retarded in much the same way that reading is taught; in this instance through instruction in the decoding of visual and auditory social cues. It was reasoned that more appropriate social behavior should accompany this increase in social comprehension. Instructional programs, described in Chapter 4, were given trial in small classes for educable retarded at the prevocational level. These programs consisted of lessons presented daily over a period varying from six to ten weeks. The lessons were prepared in the form of a teaching script illustrated by 35 mm slides, tape recorder tapes, dittoed seatwork and additional items. A sample lesson from the final version of the program is included as appendix D.

The preceding sections of this Chapter described the four experimental trials of the lessons and presented the results of each trial. The trials of the lessons provided for pre- and posttraining measurement of social cue comprehension by the Test of Social Inference (TSI), and for pre- and post pupil behavioral adequacy scores from ratings by their teachers using nine behavior rating scales (T-BR). The TSI, described in Chapter 2, yields a social inference score (TSI score) based upon pupil verbal responses to standard questions concerning pictures of social situations. The T-BR ratings, described in Chapter III, yielded a score ranging from 1 to 5 for each pupil on each of nine scales. Ratings on five of the nine scales, *social range*, *social relationships*, *peer acceptance*, *attentiveness*, and *social invisibility*, were expected to be related to pupil performance on the TSI and were included as the social behavioral scales. Ratings on *calmness*, *industry*, *academic skill*, and *appearance* were expected to be relatively unrelated to pupil performance on the TSI. These scales were included in the battery to obscure the focus of project interest. Ratings on the *social range*, *social relationship* and *peer acceptance* scales were found highly intercorrelated; hence the average of ratings on these three scales, termed the *social adequacy* rating, was reported in lieu of the ratings on each of the scales separately.

Trials of the experimental program were arranged to take place serially so the lessons and illustrations could be revised from suggestions and criticisms from teachers who used the materials in preceding trials. The initial program consisted of a set of 40 lessons and illustrative items prepared for junior high special education classes. After the first trial, the program was revised and supplemented into a total of 44 lessons for use in institutional special education classes. After two additional trials a second major revision of the materials resulted in 50 lessons intended for either public school or institutional classes or for use in other settings for the prevocational preparation of educable retarded youth. The data from the trial of this second revision, as well as information from Clark's (1967) independent replication in public school classes of the first revision, led to a third revision of the instructional materials which was published with the title *Social Perceptual Training for*

Retarded Youth (Edmonson, Leach & Leland, 1967).

The trials were arranged to compare the effectiveness of the experimental program with other educational programs, both in terms of social cue interpretation and in terms of the adequacy of the resulting social behavior. The first trial contrasted the illustrated experimental lessons with a traditional junior high special education program and with a traditional program enhanced by the addition of illustrative filmstrips and movies. In this first trial two of four junior high special education classes in a metropolitan Johnson County, Kansas, school district were randomly designated as "experimental" to receive the social cue training material. The remaining two classes, designated as "placebo," received filmstrips and movies to supplement customary units on grooming, health and transportation. A fifth class in adjacent Turner, Kansas, received the same tests and ratings but had no change made in its usual program. Trial 2 was conducted at Parsons, Kansas, State Hospital and Training Center in two sessions. In session I, two of four specially constituted classes were randomly designated as "experimental" to receive the social cue training material. The remaining two classes, designated as "contrast," were provided a second experimentally tested series of prevocational lessons which were supplemented by selected filmstrips and movies. In session II, the two programs were alternated, so that the former experimental classes were provided the contrast lessons, and the former contrast classes were provided the experimental lessons. Trial 3, conducted at Hissom Memorial Center in Oklahoma, was essentially a replication, comparing the experimental with the contrast lessons in a second institutional setting. For this trial one of two specially constituted classes was randomly designated as "experimental" to use the social cue training lessons, the remaining class being provided the "contrast" lessons. For this trial the contrast lessons were presented in their original form, without supplemental films and filmstrips. Both classes were taught by a single teacher who used the contrast program with one class each morning and the experimental program with the second class in the afternoons. Trial 4 was arranged following the evidence that boys had made greater gains than girls in the preceding three trials as measured by their TSI scores. For Trial 4 a revised set of 50 lessons was used in four junior high school education classes, two of them in a Johnson County, Kansas school district and two in the Springfield, Missouri public school district. By the time this third revision of the material had been prepared, there were not ten weeks left in the school year to allow the trial of the 50 lessons in their planned sequence. The lessons were, therefore, abridged into two series of 30 lessons. One class in each of the school districts used the lessons prepared for odd-numbered weeks (1, 3, 5, 7, 9) and included the lessons for the tenth week. A second class in each of the districts used the lessons prepared for even-numbered weeks (2, 4, 6, 8, 10) and included the lessons for the first week. In addition to comparing the TSI performance of boys and girls in Trial 4, the scores of the pupils (all of them "experimental") were compared with the scores of the control class pupils from Trial 1. Also the two abridged sets of lessons were compared with respect to associated TSI gains.

The trials demonstrated the effectiveness of the experimental social cue training lessons in producing gains in TSI scores. Both the institutional and the noninstitutional pupils, and both boys and girls, gained significantly in their interpretation of social cues as measured by the test. TSI gains by boys were greater than those by girls in the first three trials.

The contrasting set of experimental lessons used with the institutional pupils in Trials 2 and 3 also resulted in significant TSI gains, the gains being approximately the same as those made by girls using the experimental social cue training materials. The gains were less than those made by boys using the experimental lessons. The placebo audio-visual materials and the control class program in Trial 1 did not result in similar TSI gains. The results of the four trials in terms of TSI gains are presented graphically in Figure 5.1.

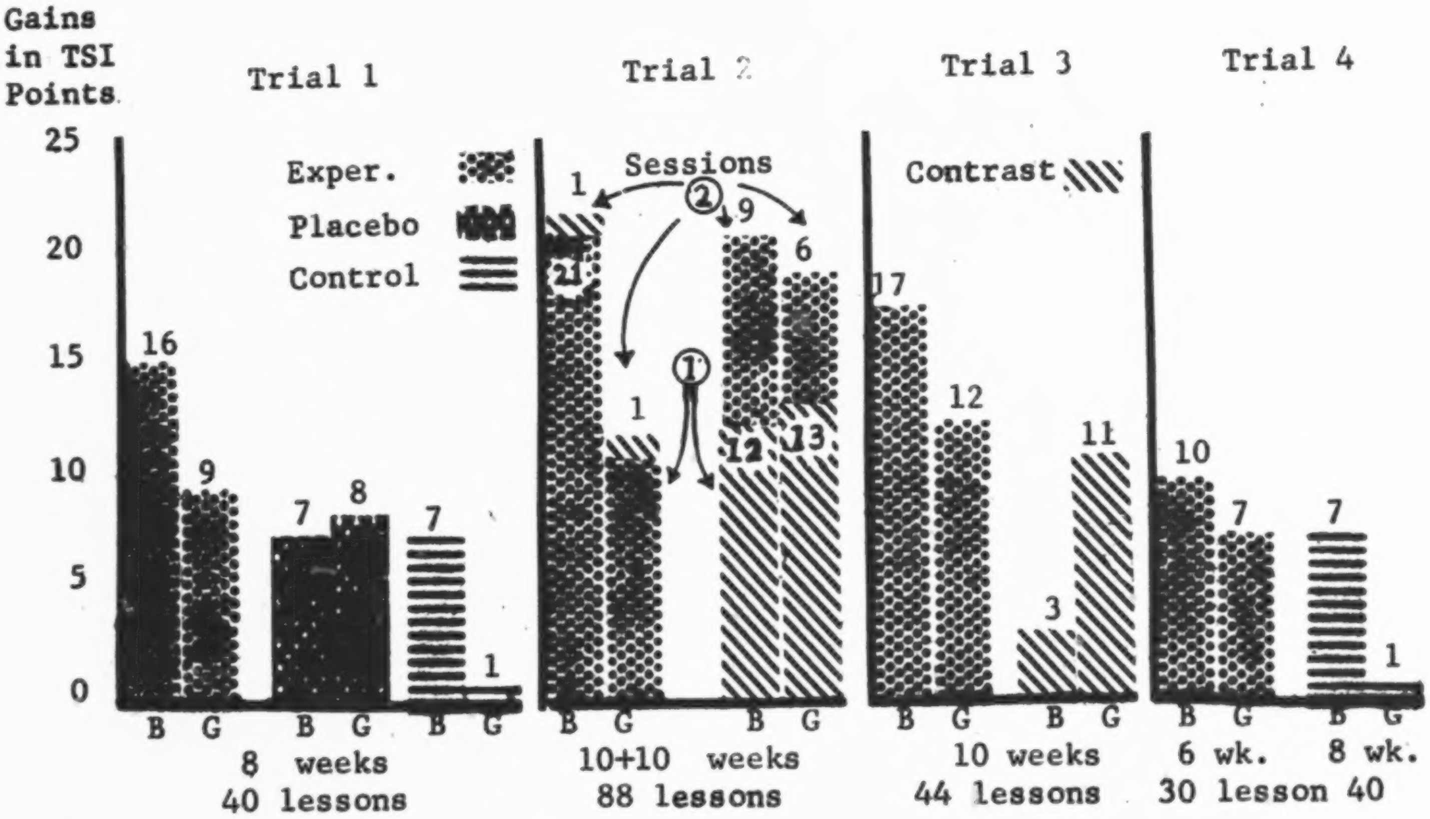


FIG. 5. TSI Gains by Pupils Using Experimental, Placebo, Control and Contrast Programs, Gains Shown Separately for Boys & Girls

In Figure 5.1 the leftmost figure represents the average gains in TSI scores, separately for boys and girls, that occurred in each of the groups in Trial 1. The boys and girls who used the experimental lessons for eight weeks had gains of +16 and +9 points respectively. The boys and girls in the placebo classes who had supplemental movies and filmstrips for eight weeks had average gains of +7 and +8 points respectively. The average gains by boys and girls in the control class who were tested and retested in an eight week interval were +7 and +1 TSI points respectively.

The next figure represents average gains in TSI scores that occurred in both sessions of Trial 2. The boys and girls who used the revised experimental lessons for 10 weeks had average gains of +21 and +11 TSI points respectively. The boys and girls who used the contrast lessons for 10 weeks had average gains of +12 and +13 TSI points respectively. The bottom portion of the four bars represent the average TSI gains in the first session. The variation in texture appearing at the top of the four bars represents the additional gains obtained after the alternation of the two curricula. Additional average gains of +9 and +6 TSI points were obtained from the boys and girls, respectively, whose use of the experimental program followed use of the contrast program. Additional average gains of +1 and +1 TSI points were obtained from the boys and girls whose use of the contrast program followed use of the experimental program.

The third figure represents average gains in TSI scores that occurred in Trial 3. The boys and girls who used the experimental lessons for 10 weeks had average gains of +17 and +12 TSI points respectively. The boys and girls who used the contrast lessons for 10 weeks had average gains of +3 and +11 TSI points respectively.

The fourth figure represents average gains in TSI in Trial 4. The boys and girls who used the abridged sets of 30 experimental lessons for six weeks had gains of +10 and +7 points respectively. These gains are shown in the figure in comparison with gains made by the Trial 1 control class pupils who were tested and retested with eight weeks intervening.

The results in terms of changes in teacher ratings did not confirm the expectation of improved social adequacy scores in relation to gains in TSI performance. Although the use of the teacher as an evaluator appeared to result in a meaningful ordering of subjects in terms of scale criteria at a given time, it appears, as discussed in Chapter III, that the gain scores that result from re-ratings after a period of time do not provide readily interpretable measures of pupil behavioral change. As was shown in Table 3.5, inverse correlations were found to exist between initial ratings and gains in ratings on almost all scales in every class, suggesting intervention of bias in the rating procedure. In response to the differing magnitudes of the correlations across teachers and across scales, a more gross nonparametric analysis of the significance of changes in T-BR ratings was used in lieu of covariance analysis. The gains were examined separately for the pupils whose initial ratings were at the level of the group median or above (initially high), and for those whose initial ratings were below the group median (initially low), on each of the rating scales. The differences thus obtained from the four trials are summarized in terms of significance at the .05 level, in Table 5.16 below.

TABLE 5.16

Summary of Significant Differences Between Groups in Four Trials in Terms of Changes in T-BR Ratings

	Significant Gains*Favoring Initially High Rated			Significant Gains*favoring Initially Low Rated		
Trial / Scales:	Soc. Adeq.	Soc. Invis.	Atten-t'ness	Soc. Adeq.	Soc. Invis.	Atten-t'ness
1	Exper > Plac.	Exper > Plac.	Exper > Plac.	none	none	none
		Exper > Control	Plac. > Control			
2 Session 1	Contrast > Exper.	none	none	none	none	none
Session 2	none	none	Boys > Girls	none	none	none
3	none	none	none	none	Contrast > Exper.	none
4	none	none	none	none	none	none

---

* $p = <.05$

As may be noted in Table 5.16, in Trial 1 the experimental pupils whose initial ratings were at the group median or above were significantly more often given higher ratings on the composite *social adequacy* scales and on *social invisibility* and *attentiveness* at the conclusion of the program than were pupils in the placebo classes. The experimental pupils received significantly more increased ratings on the *social invisibility* scale than pupils in the control class. Pupils in the placebo classes received significantly more increased ratings on *attentiveness* than pupils in the control class. No other differences in frequency of increased ratings were significant.

In session 1 of Trial 2, initially highly rated pupils using the contrast program received significantly more increased ratings on the composite *social adequacy* scale than pupils using the experimental program. No other differences were significant.

In session 2 of Trial 2, no differences in frequency of increased ratings between experimental or contrast classes were significant, but the combined group of boys received significantly more increased ratings than girls.

In Trial 3, initially low rated pupils using the contrast program received significantly more increased ratings on the social invisibility scale than pupils using the experimental program. No other differences in frequency of increased ratings were significant.

These differences may be interpreted as indicating a greater difference in effectiveness with respect to associated behavioral changes between the first experimental curriculum and the competing two programs in Trial 1 than between the second experimental curriculum and the contrast program used in both Trials 2 and 3. This interpretation is most tentative, however, because of teacher idiosyncracies in use of the rating scales.

The six teachers who used the experimental lessons in Trials 1 and 2 evaluated the lessons in terms of pupil interest, level of language and suitability of content. Most of the ratings were "above average." The four teachers in Trial 2 who used both the experimental and the contrast programs differed in their preferences for the materials. Two teachers of girl pupils in Trial 2, while rating the materials favorably, indicated their preference for the more academic approach of the contrast program. The two teachers of boys in Trial 2 indicated their strong preference for the less academic experimental material. It was concluded that both the social perceptual program and the contrast program (Fudell, 1963) are effective and appropriate for the prevocational level. It is apparent from the results of the trials that significant improvement in comprehension of social situations can be induced among educable retarded youth by instructional programs of more than one type. Gain in ratings of social behavioral adequacy were not found associated with TSI gains.

CHAPTER VI

DISCUSSION OF RESULTS AND IMPLICATIONS

I. Discussion

In this chapter certain of the results that were reported in Chapters 2, 3, and 5, concerning, respectively, the Test of Social Inference (TSI), ratings of social behavior, and trials of the social perceptual training program, are briefly reviewed and at times supplemented by additional data as a basis for a discussion of the rehabilitation implications of the project data. First discussed is the Test of Social Inference with consideration of reliability criteria and the meaningfulness of the scores. Next discussed is the social perceptual deficit of the retarded revealed by the TSI, as posing a need for rehabilitation focus. Next discussed are possibilities for remedial treatment of this deficit as demonstrated by this project's experimental programs. Next, considered, becuase of its relevance to further research, is the significant relationship noted between TSI scores and ratings of behavioral adequacy. The chapter concludes by presenting a number of implications for research as well as for remedial treatment of the retarded.

A. Test of Social Inference

It had been noted in preproject observation that in response to pictured simulation of social situations such as those illustrated in Chapter 2 (Figures 2.1 and 2.3), retarded subjects frequently failed to get the point, tending to describe what they perceived rather than making the expected inferences from the pictorial cues. Unlike responses to the pictures by NMR eighth and ninth graders whose replies were customarily focused upon socially "central" aspects of the scene, the replies by retarded subjects of similar ages included many references to socially irrelevant details or portions of the scenes.

By use of the TSI, which assigns credit of one point for each central inference and which does not assign credit for descriptive responses, for low level, improbable, or irrelevant inference,[54] differences between subjects in their performance of this social comprehension task may be indicated by their respective TSI scores.

(1) Reliability and Stability of Mean TSI Scores:

To determine the test-retest reliability data reported in Chapter 2 (pages 2.8-2.12), involved five groups of retarded subjects: three special education classes at the junior high level, one at the intermediate level, and one institutional group. For the intermediate class the time interval was one week; for the three junior high classes, the interval was nine weeks; for the group of institutional EMR there was a ten day interval. The five examiners were carefully supervised in practice use of the test. The reliability coefficients for the five classes ranged from .84 to .97. With the exception of the correlation of .84 for the intermediate class, the coefficients were .90 and above, indicating very stable ordering of pupils within these classes on two occasions with respect to TSI performance.

---

[54] See pages 2.4-2.7 (Chapter I2) for description of scoring.

Although the differentiation of performance was shown to be well maintained over the retest interval, test scores earned on the second administration were generally somewhat higher than scores earned on the initial test with 45 of the 70 pupils tested (64 percent) having higher scores on the second testing. The mean gains of the five reliability samples ranged from 1 to 7 TSI points; for three of the five groups tested, the mean gains exceeding those to be expected by chance at the .05 level. The prevalence of retest gains necessitates the use of testing control groups to assess TSI changes independent of treatment, a procedure that was followed throughout.

(2) TSI Scores in Relation to Social Behavior (Validity):

The rationale presented for the TSI in Chapter 2 clearly points to the expectation that the retardate's social behavior both good and poor should be related to the level of his social comprehension. To determine if this were so, 15 teachers were given a battery of nine graphic rating scales[55], each permitting ratings of from poor to very good along a five point continuum with respect to a particular dimension of social adequacy. The five "social" scales, social range, social relationship, peer acceptance, attentiveness, and social invisibility, were expected to be related to pupil TSI performance. The social range scale provided for the rating of pupils in respect to their ease of movement in and out of a range of social situations. The social relationship scale rated pupils with respect to their social interactions. The peer acceptance scale rated pupils as to apparent acceptability by their classmates. The attentiveness scale provided for ratings of pupil ability to focus their attention upon persons and tasks. The social invisibility scale provided for ratings in terms of the pupil's ability to fit inconspicuously into social situations. Ratings on four additional scales, calmness, industry, academic skill, and appearance, expected to be somewhat less related to pupil performance on the TSI, were included primarily to divert the rater's attention from the study focus on social behavior.

The relationships between the TSI scores and the ratings on the nine scales were examined in terms of Pearson product moment correlations. As reported in Chapter 2 in Table 2.8, the correlations between TSI scores and ratings on the scales, social range, social relationships, peer acceptance, attentiveness, and calmness, coefficients of .46, .36, .39, .26, .30, were all significant at levels of less than .001 probability. The relationship between TSI and ratings of social invisibility, industry, academic skill and appearance, although positive, did not exceed chance levels of association, coefficients of .12, .14, .14, and .08 respectively. A further evidence of a relation between pupil TSI performance and more complex social adequacy were significant (see page 2.26) correlations obtained by Clark (1967) between his EMR pupils' TSI scores and scores obtained on the Vineland Social Maturity Scale and the Fudell Test of Occupational

---

[55] The complete definition for each scale is included in Appendix B.

Readiness.

B. Social Comprehension Deficit of the Retarded as Measured by the TSI

(1) TSI Score Distributions:

This project's concern for the deficit in social comprehension of retarded adolescents and adults was amply reinforced by results from the testing of groups of public school and institutional educable retarded subjects (PS-EMR and Inst-EMR), using the Test of Social Inference (TSI). The distributions of TSI scores obtained by retarded and by nonretarded subjects were markedly different. These comparisons involved 271 retardate adolescents (163 enrolled in public school special education classes and 108 resident in state institutions) and 88 NMR eighth and ninth graders who were tested in the course of the project (see Chapter 2, Tables 2.4 through 2.7). Graphic representation of the data is included as Figure 2.2. As may be noted in Table 2.4, the range of TSI scores of the NMR pupils was from 53 to 116, with a mean TSI score of 83. The range of TSI scores of the PS-EMR pupils was from 8 to 104, with a mean of 55. The range of scores of the Inst-EMR was from 5 to 84 with a mean of 39. The means of the three distributions differed at less than the .05 level of significance.

The differences in TSI scores of the public school and institutionalized retarded subjects were further examined in terms of differing distributions of scores obtained by PS-EMR and Inst-EMR who were similar of age and IQ. These score distributions are summarized in Table 6.1 which presents the means and ranges of TSI scores for each of three age categories in relation to each of four categories of IQ. The data are presented separately for the PS and the Inst-EMR.

In Table 6.1 within each of the categories of IQ and age the mean TSI scores of the public school pupils were higher than scores of the Inst-EMR who were of the same age or older. For 10 of the 11 age-IQ categories the differences between means ranged from 11 to 30 points, the single exception being the 15-16 year olds within the IQ range of 70-79 with a difference favoring the PS subjects of only three points. The conclusion from these data is a clear handicap of the institutional retardate as compared with his same age-same IQ peer with respect to the social interpretation skills measured by the TSI.

(2) Sampling:

Before discussing the generalizable implications of the social comprehension deficit, the extent to which the particular NMR, PS-EMR and Inst-EMR groups tested are typical of other groups bearing these designations within other school districts or institutions bears more careful examination. Although it is apparent from the data in Table 6.1 that both age and IQ are related to TSI performance of the retardate, similar data for the non retardate is lacking. The required IQ data for developing similar age-IQ subgroups

Table 6.1

Means and Ranges of TSI Scores of 163 PS-EMR and 107[a] Inst-EMR Subjects in Relation to Residence, Age and IQ

IQ's	CA in Months	PS-EMR N	PS-EMR TSI	PS-EMR Range	Inst-EMR N	Inst-EMR TSI	Inst-EMR Range	diff.
59 & below	Under 179	(12)	31.4	9 to 52	(5)	18.2	5 to 28	13
	180-203	(11)	42.9	26 to 77	(16)	32.2	6 to 50	11
	Over 203	(6)	44.3	36 to 56	(19)	29.6	5 to 58	15
60-69	Under 179	(23)	49.6	26 to 90	(9)	29.7	8 to 60	20
	180-203	(25)	57.0	18 to 82	(12)	41.4	8 to 74	16
	Over 203	(3)	80.0	70 to 94	(14)	49.9	14 to 78	30
70-79	Under 179	(40)	56.0	8 to 79	(6)	38.7	28 to 45	17
	180-203	(19)	59.6	37 to 98	(6)	56.8	36 to 84	3
	Over 203	(6)	68.7	62 to 82	(12)	40.2	14 to 81	28
80+	Under 179	(9)	59.5	34 to 104	(0)	-	-	-
	180-203	(7)	76.0	68 to 84	(3)	52.0	32 to 64	24
	Over 203	(2)	85.0	84 to 86	(5)	65.6	41 to 82	19

---

[a] IQ not available for one institutional subject.

for the 88 eighth and ninth graders tested on the TSI was not available. Even if it had been, the limited age ranges and small cell frequencies would have precluded generalization. The possibility of large interclass (sample) differences for non retarded groups is suggested by the considerable intersample heterogeneity within both the public school and institutionalized retarded groups. This heterogeneity is clearly in evidence in the Table 6.2 presentation of the means and ranges of the TSI scores, IQ's, and ages of the 12 public school special education classes and the 9 institutional classes and groups for whom data is available.

For the 11 junior high school special education classes in the 12 class PS-EMR sample (the twelfth being an elementary school class) the TSI means covered a 19 point range; the IQ means covered a nine point range; and the age means covered a 39 months' range. For the nine institutional samples the TSI means ranged 15 points, the IQ means 17 points, and the age means ranged 46 months.

Table 6.2

Means and Ranges of TSI Scores, IQ's, and Ages of 12 PS-EMR Classes and 9 Inst-EMR Groups

PS-EMR Classes

Class[a]	N	$\overline{TSI}$	Range	$\overline{IQ}$	Range	CA	Range
BB	20	63	36-86	66	52-90	205	186-219
N	12	60	30-104	71	56-80	169	153-202
AA	13	60	37-94	64	51-90	200	184-229
C	15	58	36-98	67	48-77	178	159-194
A	11	55	16-84	69	51-80	180	151-199
E(CC)	12	55	27-79	70	60-79	178	147-209
L	14	55	26-73	72	62-85	179	166-195
Intermed.	15	55	38-76	69	60-79	172	141-199
M	24	53	34-78	76	52-88	166	149-183
B	8	50	9-90	64	55-71	181	171-199
D	8	44	18-74	63	57-72	172	156-192
O	11	44	8-74	70	45-98	166	152-186
Mean	14	55		69		178	
Range of Means		44-63		63-72		166-205	

Inst-EMR Classes and Groups

Groups	N	$\overline{TSI}$	Range	$\overline{IQ}$	Range	CA	Range
G	9	48	32-78	70	59-83	201	171-240
I	9	43	33-56	67	54-79	203	171-245
H	9	43	16-74	74[b]	[c]-87	198	176-217
F	9	42	14-74	72	61-84	198	179-214
#1	10	42	10-82	64	41-84	215	203-243
#2	14	38	5-77	57	46-68	234	198-243
J	17	37	8-84	61	43-77	189	148-241
#3	14	33	5-57	58	46-68	188	113-244
K	17	33	6-74	63	46-83	190	141-219
Mean	12	39		64		199	
Range of Means		33-48		54-74		188-234	

---

a Identifying code allows reference to Tables contained in Chapters 2 and 5.

b N=8. IQ not available for one subject

c One institutional pupil who scored below scale level on one or more WAIS subtests earned a DAP IQ of 88 and a Leiter IQ of 56.

A further limitation to generalization from the TSI score distributions presented in this report are the sample selection procedures themselves. All persons comprising the PS samples were members of intact classes in school districts whose participation was solicited because of their accessibility.[56] The nine institutional classes similarly were not freely selected from a pool of all possible institutionalized retardates within the desired age and IQ range, but drawn from two geographically accessible cooperating institutions.[57]

The extent to which

The extent to which the subject selection procedures constrain generalization regarding the TSI data can be only empirically decided. The similarities of TSI score distribution for Clark's (1967a) data collected independently in a Tennessee metropolitan locale lends support for argument of nonregional bias; the problem of possible distortions of, say, higher TSI scores from cooperating as contrasted with noncooperating classes is probably very minor. Across institutional TSI score differences for retardates of similar age and IQ ranges, however, may be quite large, depending perhaps, in main, upon differences in the educational and training programs offered their patients by the different institutions. Prior to some preliminary testing, the problem of generalization from the TSI scores presented in this report to any particular subpopulation of retardates is resolvable only to the limited extent that the subpopulation of interest and the subjects described in the present report may be similarly described.

C. Effectiveness of the Remedial Programs

From the project's rationale, it was assumed that the social comprehension deficit of the retarded was akin to a disability in the area of reading;

---

[56] Descriptions are provided in Chapter 5 for classes, A,B,C,D,E (page 5.2), and L,M,N,O (pages 5.24-5.25). Of the remaining three PS classes (AA, BB and Intermediate) the pupils in classes AA and BB had only recently been identified as subjects for special education. Their academic performance and social behavior in regular classes had led to their referral to newly constituted work-study programs in their respective high schools. The pupils in class BB were mostly from middle socioeconomic families living in a suburban area. The pupils in class AA were mostly from urban lower middle socioeconomic families. Pupils in the Intermediate class were from families whose occupations nand incomes resembled those of the AA sample; the neighborhood, however, was suburban.

[57] Descriptions are provided in Chapter 5 for classes F,G,H,I (pages 5.9) and J,K (5.18-5.19) The remaining Inst groups #1-#3 were groups specially constituted at Parsons State Hospital in connection with the early development of the TSI. Group 1, used as a reliability sample, were young men and women being readied for release. Group 2, all males, were a group being closely studied in a behavioral observation project (Bass, 1967). The subjects in Group 3 were selected for testing as representing subjects similar in IQ to those in Group 2, but less adequate with respect to social behavior.

or, perhaps more resembled a disability of hearing or vision. It was expected that effective treatment would consist in part of "amplifying" the visual and auditory cues, and in part of explaining what they signified. The lessons that were prepared for trial were unlike other special education materials in their de-emphasis of reading and writing and in their focus, instead, on cue decoding and cue-related social behavior. The lessons illustrated cues and provided for practice in the rudiments of daily social transactions.

As a result of trials of the successively revised editions of the remedial programs, confirmation was given to the expectation that the retardate's sizable social comprehension deficit was in large part remediable. Pre- and post TSI testing in each of the 11 classrooms receiving experimental social perceptual lessons (see Chapter 5) revealed gains in TSI scores in every instance (every class). The largest gains accompanied use in state institutions of two different instructional programs. One of them consisted of the social perceptual lessons (experimental) prepared for this purpose, the material focusing upon the signal properties of portions of the milieu, providing examples of visual and auditory "cues." The second was a program (contrast) prepared by Fudell (1963), focusing upon the behavioral expectations of employers. These two sets of lessons, although differing in content and presentation, both dramatized adult roles. Gains made by pupils exposed to these two programs were consistently greater than gains made by pupils in public school special education control classes provided audio-visual materials to supplement their more traditional program which focused in a more unitary way on such topics as health, transportation, and communication. The greater success of the two differing programs, one of them especially illustrated, the other not, points to some characteristic beside illustration as related to pupil improvement. One may speculate that the novelty of both programs, their above average interest ratings, and their focus upon adult roles, were factors in their potency. Of the 127 pupils who used the experimental social perceptual lessons, 50 percent had gains of 10 or more TSI points when tested at the conclusion of the remedial program; only 12 percent having no gains. Of the 47 pupils who used the contrast lessons, 45 percent had gains of 10 or more TSI points, 31 percent having no gains. Of the 23 pupils who used the placebo audio-visual materials, 39 percent had gains of 10 or more TSI points, 17 percent having no gains. In comparison, of the 45 pupils who had no supplement to their programs, 27 percent had gains of 10 or more TSI points and 29 percent had no gains.

The TSI gain data from the experimental classes were examined for indication that age or IQ or sex might be an important factor in remediability. A summary of the data is presented in Table 6.3.

It is apparent from examination of the table that sex was a differentiating factor in gain in association with the social perceptual lessons. For each category of age or IQ, boys' TSI gains in the experimental classes exceeded gains by girls. Counts made of individual pupil gain scores revealed that 63 percent of the boys had gains of 10 or more TSI points, whereas only 34 percent of the girls had gains of 10 or more points.[58] As has been discussed

---

[58] On the other hand, use of Fudell's Occupational Educational Units as the comparison lessons in Trials 2 and 3 resulted in greater gains by girls than boys. Only 32 percent of the boys had gains of 10 or more points, whereas 56 percent of the girls had similar gains.

Table 6.3

Mean TSI Score Gains and Ranges in Relation to Sex, Age and IQ of 125[a] Experimental Class Pupils

Category CA in months	Boys (N)	 $\overline{\text{TSI}}$ Gain	 Range	Girls (N)	 $\overline{\text{TSI}}$ Gain	 Range
179 & below	(39)	12.9	-7 to 34	(28)	6.7	-5 to 24
180 - 203	(25)	13.4	-4 to 53	(19)	8.5	-4 to 24
204+	(6)	17.8	11 to 30	(8)	14.9	4 to 26

Category IQ	Boys (N)	 $\overline{\text{TSI}}$ Gain	 Range	Girls (N)	 $\overline{\text{TSI}}$ Gain	 Range
59 & below	(8)	15.5	1 to 36	(12)	6.7	-2 to 22
60 - 69	(16)	16.7	2 to 34	(20)	8.5	-5 to 24
70 - 79	(33)	12.4	-7 to 53	(17)	10.1	-3 to 24
80+	(13)	11.5	2 to 28	(6)	7.8	-2 to 26
Total	(70)	13.6		(55)	8.5	

---

[a] IQ's not available for two experimental subjects

elsewhere, the consistently greater gains by boys was believed due to the greater appeal to boys than to girls of the first two sets of experimental materials. Trial of the third set of social perceptual materials, especially rewritten to provide for increased female relevance, resulted in more equal gains by the two sexes. The earlier finding of differential gains, however, demonstrates the value of field tests of remedial techniques.

Examination of the average gains of experimental class pupils in relation to age, reveals consistently larger average gains for each increasing category of age. This relationship is less apparent for boys than for girls. The average gain by girls 15 years old and younger was less than half the average gain of the small sample of girls 17 years old and older.

Examination of gains in relation to IQ reveals a different pattern. The relationship for boys is negative, the highest average gain being associated with the next lowest IQ category, and the lowest average gain being associated with the highest IQ category. The relationship for girls, although positive, is less pronounced than the relationship between TSI gain and age.

Examination of the columns which present the ranges of TSI gains, reveals that high gains were obtainable within each category of age or IQ. From the data it does not appear that the remedial effect of the experimental lessons is importantly limited by either age or IQ, within the span of ages and IQ's of the subjects here represented.

Factors other than age and IQ are apparently related to TSI gains. Because of the sample size, it appeared impractical to make further separations of data within Table 6.3; however, there were overall differences in gains between PS and Inst subjects, the average gain by the Inst-EMR exceeding those by the PS subjects for most categories. The difference was particularly apparent in the case of boys. The overall mean gain of 10 TSI points of the PS-EMR boys was exceeded by the overall average gain of 16 TSI points of the Inst boys. The difference was less apparent in the case of girls. The overall mean gain of the PS girls of 7.6 TSI points was slightly exceeded by the mean gain of 10.2 TSI points of the Inst girls.

An additional variable in gain appears to be teacher experience, perhaps more accurately termed skillful handling of pupils and materials. Among the experimental classes, the smallest overall gains were exhibited by pupils in two classes taught by teachers completing their first year of teaching.

The gains that were obtained during the brief experimental use of the lessons should not be regarded in any sense as maximal. The teachers who used the experimental lessons received no orientation, being merely provided a kit of "canned" materials and a teaching guide whose sequence had to be followed and completed within a stipulated period of time. Teachers had no choice of materials, and it became apparent from the teacher evaluations of the lessons, in Trial 2, that teachers have preferences! (See Chapter 5, pages 5.17-5.18) In spite of such limiting factors, the gains in TSI scores by some individuals placed them well within the range of TSI performance by NMR pupils. In terms of TSI scores, their social comprehension was the equivalent of the NMR.

A question related to remediability concerns the permanence of the gains in social inference performance. Although it might be expected that increased comprehension would be reinforcing to visual attentiveness, information in this regard is lacking since the posttests were conducted no longer than a week after completion of the experimental program. A substantive answer to the question of stability of change must await a follow-up study of experimental pupils.

D. Relation of TSI Gains to Gains in Behavioral Ratings

Although, as has been shown, a relationship exists between TSI performance and teacher ratings of certain aspects of social functioning, gains in TSI scores associated with the special instructional treatment were not correlated with gains in these ratings. Product moment correlations between TSI gains and teacher rating gain on the social range scale ranged for nine[59] classes from -.24 to .31, with a median r of .04. The correlations between TSI gains and TR gains on the peer acceptance scale ranged from a low of -.36 to a high of .29 with a median r of -.12. The correlations between TSI gains and TR gains on the social invisibility scale ranged from a low of -.38 to a high of .48 with a median of -.17. These low indices of correlation may, of course, be truly indicative of a lack of functional relationship between increased comprehension and more adequate behavior. However, in addition to general considerations of statistical unreliability of change (Harris, 1963), there are factors indigenous both to the test and to the teacher as a rater that reduce the possibility that the two sets of gain scores would vary for an individual in the same direction and in the same proportion. Considering the test factor first, there is the problem of equivalent gain score units. In scoring the TSI items each good inference was given credit, the number of points earned on a single test picture varying from 0 to as many as 7 points. Since the test pictures vary in difficulty, but scores are not weighted, one subject's posttest response to an easier picture of, say five more creditable inferences than he made on his pretest would be counted as the equivalent of the gain in comprehension by another subject whose posttest increase of five new inferences was related to more difficult pictures. To the extent that the foregoing occurs, the "summary gain score" computed from the TSI data is necessarily inaccurate as a measure of increased social inferential ability.

---

[59] The correlations are omitted for the two institutional classes (N's of 6 each) whose pupils initially used the contrast materials.

With respect to the teacher as a rater, examination of the ratings revealed that rating changes were in most instances inversely correlated with initial ratings. The surprising prevalence of this inverse relationship is shown by the listings in Table 3.5 in Chapter 3 of the 128 correlations between initial ratings and changes in ratings. Ninety-eight of these correlations were negative and two-thirds of the negative correlations exceeded -.50. To some extent this trend may result from the five point ceiling on the rating scales which make it impossible for pupils whose initial ratings were at the top to gain in rating. It is also conceivable that some teachers were reluctant to rate many of their retarded pupils at the top of a scale. It is speculated further as an explanation for the preponderantly negative relation, that teachers may especially desire to see progress by the socially lagging pupil; hence may be especially attentive and more noticing of his gains to the neglect of the more adequate pupil who may also have gained. Whatever the case, the preponderance of negative correlation between gain and initial rating casts serious doubt upon the validity of rating change scores. Although a teacher's rating does appear to provide a meaningful differentiation between pupils when first rated, the shift in ratings over time appears less interpretable. If one is to assume that change scores are real reflections of changes in social adequacy, one must accept that in the majority of cases initially inadequate pupils become more adequate while initially more adequate pupils become less adequate. More persuasive to the present investigators is the conclusion that teacher rating gains do not provide a valid measure of behavioral change.

II. Implications

A. Implications for Research and Rehabilitation of the Social Comprehension Deficit

(1) Need for Remediation:

Although the widespread and significant social comprehension deficit of the retarded was not unexpected, it was a surprise, by use of the TSI, to find a wide range of TSI performance at each level of IQ and age. In the public school classes there were a few pupils whose social comprehension was apparently the equal of nonretarded pupils of the same age. At the other extreme, there were so-called EMR pupils in public schools whose lack of social comprehension appeared almost total,[60] who could make only partial interpretations of the simplest of the test items. Such pupils with extremely low scores were in some ways less behaviorally conspicuous or annoying in their classrooms than others with greater comprehension (higher TSI scores). Because of such pupils' somewhat passive conformity to instructions, their capacity to function independently could be easily overestimated. Their TSI responses, however, revealed their inability to grasp the implications of some of the simplest of cues that are normally used as guides to the self-modulation of behavior. Such low scoring persons appear pressingly in need of remedial training in social comprehension. Similarly extreme variations were found among institutional residents, but with an overall greater deficit as was illustrated by Tables 2.4 to 2.7 and by Figure 2.2 in Chapter 2.

(2) Cause:

These discoveries arouse questions as to cause. What could account for the generality of the deficit and at the same time account for the wide ranges in competency? What could account for the greater deficit of the institutional subject? These questions, though extremely relevant to prevention and rehabilitation are, unfortunately, not answerable by the current project. At the most, the present data provides a basis for conjecture. Three hypotheses pointing in turn to personality differences, to differences in parental management, and to differences in social opportunitites, are suggested:

The hypothesis of personality factors accounting for the social interpretation deficit is suggested by the TSI and teacher rating correlations (see Table 2.8). The highest correlation between teacher ratings and TSI scores was that with the scale social range, measuring the pupils' ease of movement in and out of a range of social situations. Almost equally high correlations were obtained for correlations of TSI scores with ratings of social relationship and peer acceptance. On the other hand, there was a low correlation between TSI performance and ratings on industriousness, academic skill, and social invisibility. This division of relationship suggests the high TSI performer as being more turned on by social activity than by academic tasks. He is not socially inconspicuous; therefore, he is perhaps less vulnerable than the low TSI scorer to environmental rebuff. Although there might be value in a study of personality correlates, it is satisfying to note a generality of the TSI gains by the experimental pupils indicating that remediation is not overconstrained by personality differences.

---

[60] Examples of TSI responses by several such pupils are presented on page 2.18.

The second hypothesis that the wide range in social comprehension results from different quantities and kinds of parental instruction derives from data reported in numerous studies of parent-child interaction such as the Fels studies (Baldwin, et al, 1945; 1949) and the studies by Hess and Shipman (1965) which consensually demonstrated higher levels of functioning by children whose parents used reasons or instructive explanation rather than unexplained command. In this connection, however, it is well to recall the operant conditioner's reminder of the potency of behavioral consequence in forming behavior (Lindsley, 1964; 1966). The merely instructive parent would be expected to be less effective than the parent who arranged pleasurable or unpleasurable consequences as appropriate sequel to his childrens' behavior.

To extend the speculation of the influence of environmental difference, is it possible that the diagnosis of retardation may have an "itrogenetic" or diagnosis-treatment effect? How often, we wonder, do parents respond to the diagnosis of "retardation" by a diminution of instruction and an unrealistic lowering of behavioral expectations? To the extent that this may occur, the retardate's social development would lag behind that of another who was expected to perform socially at a more average rate. In connection with rehabilitation research, it would be worthwhile to study the patterns of parental expectations, communication and guidance in relation to retardate adequacy. Meanwhile, parents of the retarded may need encouragement of their efforts to shape social behavior and might profit from "how to do it" workshops such as those provided parents of children who are deaf.

The third hypothesis is that the obtained differences in pupils' social comprehension are a result of differences in range and quantity of exposure to social situations. The Kansas ecological studies, notably those by Barker and Gump (1964), reported an inverse relationship between size of school and community and the number of social settings in which marginal pupils play participating roles. This would suggest that pupils living in smaller towns and attending smaller schools would have higher TSI scores, and that pupils living in larger towns and attending' larger schools would have lower TSI scores. In the current study, a comparison between TSI scores of pupils in schools of various size cannot be fairly made because of other possibly important differences in environment. For instance, whereas all of Barker's subjects lived in well established, older, communities; in the present study only those pupils in the smaller schools lived in comparably well established areas; the pupils in the larger schools residing in suburban areas with contingent problems of pedestrian mobility.

Community based factors other than size or "age" of community are additionally relevant. In any study of retarded pupils, the quality and extent of the school's program is a crucial variable. The lower range of TSI scores of two of this project's samples might be accounted for by the lack of facilities for retarded children with greater deficits; that is, the absence of special programs for trainable retarded in these particular larger schools. The preponderantly lower TSI performance of the EMR in institutions is similarly clouded by confounding interpretations.

Although there is plausibility to the argument that the greater deficit is caused by their lesser need for social awareness (because of the prosthetic nature of their surroundings), and the lesser operation of natural consequences in cases of unawareness, there is likewise plausibility to the argument that an initially lower social comprehension (due to home environmental conditions or disturbed personality factors)led to institutional placement. A "causative" explanation of social perceptual deficits requires accompanying descriptive social ecological studies, such as those by Barker, Wright and Gump, to reveal the specific settings in which retarded subjects function, the conditions for their social successes and failures, and, perhaps most important to the rehabilitation worker a more precise delineation of their behaviors in the successful and unsuccessful conditions--actual examples of what is specifically adaptive and maladaptive. Schellhaas (1968(a); 1968 (b)) is attempting to sample community tolerances by a survey technique. Additional studies are needed--studies charting the retardates' differential responsiveness to these tolerances, his learning,and failure to learn,to cope with these situational demands.

B. Implications for Rehabilitation and Research of the Remedial Gains in Social Comprehension

Important implications of the gains that accompanied use of the experimental lessons are that the social comprehension deficit of the EMR is largely remediable and that a "canned" educative approach can be effective. Although differences in effectiveness of teacher use of the materials were noted (seemingly in part related to differences in experience and liking for the approach),through the device of the teaching script gains were obtained from nine out of ten of the 127 experimental subjects who were taught by 11 different teachers. The availability of the lessons and illustrative materials in the form of a kit raises the possibility that persons other than classroom teachers--ward personnel, counsellors, vocational trainers, volunteers, and parents--could be effectively employed in teaching "social comprehension".

The generality of gains after a relatively short period of exposure to the special lessons raises the possibility of additional gains obtainable with additional effort. Studies of the effects of extending lessons over a longer period of time and the effect of lesson supplement might be fruitful through indicating a most efficient combination of educative activities and their duration.[61]

A further question concerns an optimal age for either remedial or preventive measures in this area. From this project's data it appears that age had some effect upon the rapidity of gain. However, within the population studied, pupils as young as 12 responded with appreciable improvement. Pupil interest is, no doubt, an important factor here. Because this project's special materials were written with adolescent interests in mind, it would be expected that these specific lessons should have less appeal to much younger or to much older subjects. There seems no reason, however, why suitably prepared illustrated lessons should not be effective with retarded at the preschool or primary level, nor, indeed, at the adult level.

---

[61] In the published edition of the social perceptual lessons, suggestions for supplemental activities are included with each lesson.

An additional question, and an important one, concerns the permanence of TSI gain. Because no followup of TSI testing occurred in the present project, the stability of the changes cannot be shown. It seems most probable from the evidence of behavioral research that if an increase in comprehension were not functional for a subject--were not useful or satisfying--that comprehension and visual attentiveness would not be maintained. It has been suggested that the ability to cope with an 'expanding universe' is the chief source of reward for both normal and retarded, that any real increase in comprehension is by definition 'functional' and rewarding. To what extent and under what conditions this"self paying" system endures is answerable only in terms of more extended studies of former experimental pupils.

The further and more crucial question concerns the extent to which increased attentiveness and understanding of visual cues presented in the form of pictured material is generalized to the live social situation. Evidence is currently lacking for the expected relation between increased social comprehension and more adequate behavior as measured respectively by TSI score gains and gains in teacher ratings. However, there is good reason to question the meaningfulness of the teacher rating gains and there is no reason to assume that the present method of scoring the TSI, by totalling inferences on the basis of equal credit for each, is the best way of quantifying social comprehension. An obvious implication of the impasse is the need for further analyses and a fresh set of data. In making further anlyses of the current TSI data, it is conceivable that a comparison of responses of the more socially adequate with responses of the less adequate would reveal that a change in scoring from simple numerical addition to the differential weighting of responses would lead to the greater validity of score gains. As a source of helpful fresh data, a series of validation trials would provide new criteria for use in item analysis, with the goal of determining whether particular types of response to certain test items are predictably associated with the retardate's more successful transition to self maintenance.

Meanwhile, the hypothesis that social comprehension is important to the retarded can be more firmly postulated on the basis of the relationships, other than the gain score, demonstrated in this project's data, and it seems a reasonable conclusion to advocate the use of preventive and remedial techniques.

VII. OVERVIEW

This chapter, providing an overview of the project activities, recapitulates the rationale, goals and procedural design; describes the test devised as a measure of social comprehension; describes rating scales used as a measure of behavioral adequacy; describes lessons developed for social comprehension training; the experimental trials of the lessons; then, summarizes the results and implications.

A. Rationale, Goals, Procedural Design

Numerous attempts have been made to determine variables relevant to the postinstitutional, postschool, postrehabilitation, success of the retarded. Manual dexterity, attentiveness, lack of physical defect, IQ, and age have at times been found associated with successful adjustment.[62] At times, as cited in the opening chapter of this report, investigators have concluded that aspects of social competence not well defined nor measured by existing instruments seemed to be the more crucial factors. One reason for this interest has been the desire to predict that certain individuals lacked job-requisite skill or that they might not sufficiently profit from particular programs to warrant their admission. The interest of the present investigators in factors identifiably associated with success is, however, related to the more general rehabilitation problem of determining what is crucial to independent living so that it may be emphasized or compensated for in the retardate's preparation.

This project's concern has been the failure of the retarded adolescent to comprehend simple social situations. In a pilot study predating the present project, pictures of social situations were presented for interpretation to retarded and nonretarded subjects of varying ages. Analyses of the responses revealed that the nonretarded adolescents tended to restrict their reports to socially relevant portions of the scenes, getting to the heart of the matter while omitting reference to irrelevant and peripheral details. In contrast, the retarded tended to describe or enumerate details, stopping short of their implications. When they grasped the implications of some portion of a scene, this portion was frequently not the most relevant. This interpretative deficit of the retarded has been observed by others as is indicated in the literature on clinical testing (Beck, 1952; Beier, Gorlow & Stacy, 1951; Butler, 1961; Denhof & Robinault, 1962; Gorlow, Butler, & Guthrie, 1963); however, the prevalence of the deficit and whether or to what extent it might yield to remedial treatment has previously apparently not been investigated.

Reasoning that the retardate's lack of social comprehension contributed to his frequent behavioral deviance (thus to his rehabilitation nonsuccess) the project hoped to demonstrate this relation. The procedural stages consisted of developing a test of social comprehension and a battery of behavioral scales by which to quantify the two variables. Next developed was an experimental program by which to attempt to increase the retardate's social comprehension while observing the effect upon the behavioral scales. Project activities relating to these objectives are reviewed in the following sections.

---

[62] In a recent study Chaffin (1968) presented evidence that rate of productiveness was a significant factor in the employability of the EMR.

B. The Test of Social Inference:

Development: The Test of Social Inference (TSI) was developed through a series of pilot trials involving small groups of retarded and nonretarded subjects. In these trials, pictures of social situations, accompanied by questions as to their meaning, were used to stimulate verbal responses. Through these trials ambiguous items were eliminated and a technique of administration and method of scoring developed. The final experimental test (TSI) consisted of 36 pictures each mounted on heavy approximately 9x 11 cardboard pages to be shown one at a time to the subjects in individual test sessions. A check list recording form was developed to simplify the recording of the most frequent responses. To facilitate scoring, a scoring guide listing the several categories of responses and their scores was also prepared. Subject responses were scored by crediting each good inference that pertains to a central portion of the pictured activity and giving zero credit for merely descriptive or irrelevant portions of a response. The subject's total TSI score was defined as the sum of his inference score credits earned in a given test administration.

Reliability: Prior to use of the test as a measure of the remedial effect of the experimental social perceptual lessons, the test reliability was examined in terms of Pearson product moment correlations computed from scores obtained from two administrations of the test to four classes of special education pupils (PS-EMR) and to a group of institutional residents (Inst-EMR). The retest intervals were one week or nine weeks. The five test-retest correlations ranged from .84 to .97, indicating that test scores reliably differentiated EMR subjects over the retest intervals investigated. Retest scores were higher on the average by from 1 to 7 TSI points than scores earned on the initial presentation, the retest gains by three classes exceeding the level of chance indicating a need for nontreatment, control groups in studies in which the test may be used to assess gains in relation to treatment.

C. Social Behavior Rating Scales:

As the best available source of behavioral information, the teacher was used as the behavior rater in each of the treatment trials. In addition to the teacher as a rater, in connection with trial 3 of the experimental program at one state institution, six persons, the chaplain, nurse, librarian, school secretary, director of cottage life, and recreation director, were used as raters of outside school behavior.

Because of the project interest in social performance over a broad range of settings, nine definitions of behavioral adequacy were developed which appeared relevant to school or postschool society, each definition, in effect, constituting a criterial scale. Five of the scales were expected to be related to pupil TSI performance. These five, termed the "social scales," were the following: social range (pertaining to ease of movement in and out of diverse social settings), social relationships, peer acceptance, social invisibility (pertaining to the subject's ability to fit inconspicuously into social situations), and attentiveness. Four additional scales, included primarily to obscure for the rater the project interest in social behavior, were not expected to relate closely to TSI scores. These four scales were the following: calmness, industry, academic skill and appearance. This battery of T-BR scales provided for ratings of a behavioral dimension in terms of five categories of from "very poor" to "very good", each category being illustrated by a brief description.[63] The rater indicated the matching of pupil behavior with rating category by a check mark in the appropriate column on the form. The outside rater scales (O-BR) retained the five point graphic format and categorical description of the T-BR, but included use of explicit frame of reference procedures resembling those modified by de Jung and Haring (1962) from those used with the Syracuse Scales of Social Relations (Gardner & Thompson, 1956). In the O-BR scales an additional modification was the merging of the T-BR social relationship and peer acceptance scales (the ratings having been found highly inter correlated) into a single scale of social acceptance. The scale thus eliminated was replaced by a scale for the rating of honesty.[64]

D. Social Perceptual Training

The instructional program evolved via the sequence of preparation of experimental materials, classroom trial, revision based upon evaluation or suggestions from the using teachers. The search among prior investigations of social perception for useful tactics or research results revealed that social perception research, as generated by early research interest in empathy, had focused mainly upon a narrow band of the spectrum of social signals--those relevant to human emotion. A number of studies have investigated the interpretability of facial expressions (Harrison, 1964; Orr, 1960; Osgood, 1966; Secord, 1958). A few have examined the interpretability of posture and gesture (Ekman, 1965; Rosenburg & Langer, 1965). The present authors, desirous of teaching the retarded to recognize types of situations in relation to culturally expected behaviors, viewed "social perception" as the decoding of any of the kinds of signals that may function as cues to expected behavior in social settings.

As the remedial approach that was chosen, lessons were prepared for classroom use for pupils at the prevocational level, lessons being drafted in the form of a teaching script. In all, three lesson series (Curricula I, II and III), differing to some extent in content and presentation, were prepared and given trial sequentially in the course of the project. Curriculum IV (Edmonson, Leach & Leland, 1967), a product of successive revisions, is a final version.

---

[63] T-BR scale definitions and illustration of format are presented in Appendix B.

[64] O-BR scale definitions and illustration of format are presented in Appendix C.

The lessons of each series began by calling attention to the variety of signals that are used to direct and regulate behavior; as an example, the use of colored lights, geometric shapes, bells and horns such as might be encountered at a railway crossing; or, as another example, the gestural languages of the sports official and the Navy deck officer. As the lessons progressed, a number of adult settings important to the retardate's independent functioning were illustrated. Each lesson attempted to clarify certain cues by which to identify expected patterns of behavior. The illustrations were chiefly photographic slides and recorder tapes.

Throughout several lessons a general topic or unit was presented. Each lesson included a statement of the main objective and a list of the lesson's vocabulary which might require teacher explanation. Each lesson was introduced by a prepared statement made by the teacher followed by the first slide. The slide illustrated the point of discussion and was used as a focal point for pupil response. Lessons were prepared for verbal and visual presentation without requirement that pupils read or write. As part of the format, a verbal quiz followed every few lessons. Pupil verbal responses to the quiz were to be tallied as correct or incorrect. Correct responses were to be rewarded. All of the illustrative materials: slides, tapes, dittoed seatwork, driver handbooks, merchandise cards, and miscellaneous articles such as tape measures, playing cards and mail order catalogs required for use with certain lessons were provided teachers in kit form. The pupils' incentive awards were also provided.

E. Experimental Trials:

To examine the effectiveness of the social cue training lessons, four trials were conducted in which one or another of the three versions of the experimental lessons were used in classrooms. In all, 11 teachers used the materials during the formal trials that were a part of this project[65]. In three of the four trial situations, comparison treatment groups were used to control for factors of experimental involvement (Hawthorne effect) and unreliability of criterion measures. Table 7.1 presents a summary description of the four trials. For each trial the population (public school, special education, or institutional) is indicated in the first column of the table. The designation of the sample groups (experimental, placebo, contrast, control) and the number of boys and girls in each group are indicated in the second and third columns, respectively. The fourth column with the heading "Treatment" indicates which of the three versions of the social perceptual materials was used for the experimental treatment, and what special materials were used

---

[65] In Clark's independent trial (1967) three additional teachers used and evaluated Curriculum II.

TABLE 7.1

Summary of Trials of Social Perceptual Curriculi

Trial	Population	Groups	M	F	Treatment	Assessment Criteria
1	EMR-Junior High Special Education	2 Exper.	12	7	Curriculum I - 40 lessons, 8 weeks	Test of Social Inference (TSI)
		2 Placebo	14	9	Filmstrips & Movies 2-3 times weekly	Teacher Behavior Rating (T-BR)
		1 Control	6	6	Usual Program	Tr. Evaluation of materials
2	EMR-Inst Residents at Prevocational Level	4 Exper.[a]	15[a]	15[a]	Curriculum II - 44 lessons, 10 weeks	As above
		4 Contrast	15[a]	15[a]	Occupational Educational Units, supplemented by filmstrips & movies	
3	As above	1 Exper.	7	10	Curriculum II - 44 lessons, 10 weeks	TSI, T-BR Outside classroom Behavior Rating (O-BR)
		1 Contrast	7	10	Occupational Educ. Units w/o visual supplement	Tape recorded Tr. suggestions
4	EMR-Junior High Special Education	4 Exper.	37	24	Curriculum III - Total of 50 lessons divided into 30 lesson portions. Two classes used "odd numbered" lessons; two classes used "even numbered" lessons.	TSI, T-BR, Tr. suggestions.

---

[a] In Trial 2 two training sessions were held. The first session involved a comparison between two groups of 18 pupils (9 boys and 9 girls each) of the social perceptual curriculum II with the equally experimental Occupational Educational Units. After completion of the first trial and after a rest of approximately two months during which several pupils were placed on jobs or released from the institution, the remaining pupils from the original groups used the opposite treatment. For statistical purposes group sizes were equalized by random elimination of pupils in session two with resulting N of 6 boys and 6 girls completing the experimental and the same number and division of sexes completing the contrast treatment.

by the comparison group; the duration of the treatment periods is also indicated in this column. The trials of the experimental lessons covered periods of from six to ten weeks, lessons taking from approximately an hour (Curricula I and II) to two hours per day (Curriculum III). The assessment column lists criteria used in connection with each trial.

All pupils received pre- and posttraining testing by the TSI and pre- and posttraining behavior ratings made by their teachers on the nine rating scales. Teacher evaluations or suggestions were solicited following each trial as a means of improving the training program.

In Trial 1 the first series of lessons, Curriculum I, was tried out in two public school junior high school special education classes. Two similar (placebo) classes were provided audio-visual materials to supplement educational units on health, personal appearance and transportation. An additional ("no treatment") class had no alteration made in its customary program. The first series of lessons were modified, based in part upon teacher evaluation, into Curriculum II, which was tried out in Trials 2 and 3 in institutional classrooms.

Trial 2 involved two classes of boys and two classes of girls at Parsons State Hospital and Training Center at Parsons, Kansas. The four classes were made up from older patients in the upper adaptive behavior levels (mostly AB I or II) who were at that time in the prevocational education program at the hospital. To create the four classes, subjects were roughly ordered by sex groupings according to age and IQ and alternately assigned into one of two same-sex groups of nine subjects each. The groups were assigned by lot to a same sex teacher; then the teacher and his class were assigned by lot their curriculum treatment. Since teachers were apt to be aware of procedures in other classrooms, all four classes received special "experimental" curricula. In a first session, one class of boys and one class of girls used the social perception Curriculum II for a ten week period while the second class of boys and the second class of girls used the Occupational Educational Units (Fudell, 1963) by way of contrast. After a lapse of approximately two months the classes were reconvened for a second 10 week session in which the former experimental pupils used the contrast materials and the former control pupils used the social perception curriculum. In session 1, pre- and posttreatment TSI scores were obtained from the 36 pupils. Teachers rated pupils on the nine T-BR scales weekly, commencing with the first week of the lessons and continuing throughout the 10 week treatment. For session 2, the second TSI testing (the session 1 posttest) provided the pretreatment TSI score. A third administration of the TSI followed completion of the second session trial. Teachers began their T-BR behavior ratings the first week of the second session and made weekly ratings throughout the 10 weeks. Teachers during both sessions were requested to make daily evaluations of their sets of lessons with respect to interest, language, content, presentation, and pupil assimilation.

Concurrent with the second session of Trial 2, an additional trial, Trial 3, of the same contrasting sets of lessons was underway at a second state institution for the retarded, Hissom Memorial Center at Sand Springs, Oklahoma. Trial 3 involved two classes of pupils approximately a year younger than the Parsons subjects and somewhat lower of IQ. The available pupils were in two classes, an older and a younger group. To create two more comparable classes, these pupils were regrouped into same-sex pairs roughly equated with respect to age and IQ. Two extreme subjects in terms of age were dropped from the sample. One pupil from each pair was then reassigned by lot to one of two classes of seven boys and ten girls. One class was designed by lot to receive the social perception program (experimental). The second class received the Occupational Educational Units (contrast). Both classes were taught by the same teacher. The contrast class met with her in the morning; the experimental class in the afternoon. Lessons for both classes were used daily for 10 weeks. As in previous trials, pre- and posttreatment TSI scores were obtained from each pupil in individual testing sessions and the T-BR form was used to provide pre- and posttraining ratings of behavior. In addition an "outside school" rating of behavior (O-BR) was obtained from six nonteacher members of the institution staff. The teacher tape-recorded her criticisms and suggestions for improvement of the two sets of experimental materials.

The fourth trial of social perceptual training had not been anticipated in the original project proposal. However, the results of the preceding trials had indicated that although the experimental lessons were associated with significant improvement in social cue interpretation, the major gains in each instance were made by boys. Girls in the experimental classes, although showing sizable TSI gains, made gains no greater than those by girls taught by the contrast curriculum. A project extension was granted to permit an additional revision and trial of the experimental lessons with the intent of making them more interesting to girls. The lessons were revised for Curriculum III based upon the suggestions and criticisms of consultants and teachers who had used the lessons. Among the changes, the vocational setting lessons contained in Curriculum II were replaced with new lessons affording a broader "social mobility" orientation, with emphasis on social interaction and leisure time activities. Although teachers and pupils had responded favorably to "vocational" lessons included in Curriculum II, these lessons were replaced so as to supplement or complement Fudell's contrast curriculum rather than to overlap it. The revised lessons used less repetitious, higher level language than the previous revisions. The limit of one lesson a day for ten weeks was retained but each lesson was extended in length.

The Trial 4 of this revised program involved four classes of junior high special education pupils, two classes in the Shawnee Mission, Kansas, district, and two in the Springfield, Missouri, school district. One of the Springfield classes, normally an "8th grade", was enlarged for the experimental period by the inclusion of a "7th grade" class as a means of enlarging the sample of girl pupils. This combination class met daily for approximately two hours with the 8th grade special education teacher, who was assisted during this period by the 7th grade teacher. The two Shawnee Mission teachers were inexperienced, both completing a first year. The Springfield teachers were experienced, although the teacher of one group was serving her first year in a special education class.

The testing of the revised curriculum ran into a serious timing problem. By the time the revision had been completed, the school year was too near its end to allow the sequential trial of the ten week series of 50 lessons. Only six usable weeks remained. To permit trial of all of the lessons, although at the cost of disrupting lesson sequence and the repetition of concepts, the revised curriculum was shortened to two partially overlapping "halves," one half consisting of lessons for weeks 1,2,4,6,8 and 10 ("even numbered") and the second half consisting of lessons for weeks 1,3,5,7,9 and 10 ("odd numbered"). One teacher in each district was selected by lot to receive the "even numbered" lessons; the "odd numbered" lessons going to the second teacher in the district. As in previous trials, pre- and posttreatment TSI scores were obtained and the T-BR form was used to provide pre- and posttraining ratings of behavior.

Altogether, during the two and a half years of trials of the experimental lessons, including the comparison classes, a total of 15 special education classes, 185 subjects (101 males and 84 females), were involved in the comparison and evaluation of treatment.

F. Results and Implications:

Relationship between the Retardate's Social Comprehension Deficit and His Behavioral Adequacy.

At the opening of this chapter, in connection with the investigators' interest in the preparation of the retarded for postschool or postinstitutional living, the objective of learning whether the retardate's poor comprehension of social cues is related to his social behavioral inadequacy was indicated as a project goal. With this in view, the correlations between pupils' TSI scores and the ratings assigned them on the nine T-BR scales (with special interest focusing upon the five "social scales") were examined for the 115 PS-EMR and 70 Inst-EMR pupils involved in the four trials. In support of the anticipated relationship, the correlations between TSI scores and ratings on four of the five "social scales" (social range, social relationships, peer acceptance, and attentiveness) were found significant at less than the .001 level. Ratings on the calmness scale were found correlated at the same level of probability. Ratings on the social invisibility scale, however, the fifth "social scale," were not significantly correlated with TSI scores. It was noted from the intercorrelations of the ratings on the nine scales that social invisibility ratings by teachers are somewhat more influenced by pupil industry, academic skill and appearance than by the TSI-related characteristic of being at ease in social situations.

A generally similar pattern of relationships was revealed between the tests and ratings of the 34 Inst-EMR subjects by six nonteacher members of the institutional staff, using the nine O-BR scales. The correlations between pupils' TSI scores and O-BR ratings on social range and social adequacy (the combined social relationship / peer acceptance scale) were significant at less than the.001 level. Correlations between TSI scores and O-BR ratings of calmness and academic skill both attained significance at the .05 probability level.

Additional data in support of the hypothesized relation between comprehension and behavioral adequacy was noted in the differing distribution of TSI scores obtained from the NMR, the PS-EMR and the Inst-EMR subjects tested. From the test rationale it would be anticipated that the TSI scores of nonretarded (NMR) pupils should exceed those of retarded pupils; that the TSI scores of low IQ pupils tolerated within the regular class should exceed those of special class pupils (PS-EMR), and that the TSI scores of retarded subjects in the public school special class should exceed those of retardates whose behavioral nonadaptation had led to their institutional placement (Inst-EMR). In each instance these expectations were supported by the project data. Within each category of IQ and age of the retarded subjects tested, the average TSI score of the PS-EMR was higher than that of the Inst pupil. The TSI score distributions of the NMR, the PS-EMR and the Inst-EMR, although overlapping were reliably different. Of the NMR, only three percent scored as low as the mean score of the PS-EMR, and only six percent of the latter scored as high as the NMR mean. Of the Inst-EMR sample, fewer than 20 percent scored as high as the PS-EMR mean. Only three percent of the Inst-EMR pupils scored as high as the NMR mean. In part the greater overlap of the TSI score distributions of the PS-EMR and the Inst-EMR samples is attributable to the selectivity of the institutional samples. With the exception of one group of 14 subjects, the institutional subjects tested were top level; functioning at the prevocational level in two well-rated institutions.

With respect to a TSI score differentiation between pupils of low IQ enrolled in regular classes and those enrolled in special education, no specific search was made for a sample of the former. Among one of the samples of supposedly nonretarded eighth and ninth graders, however, eight pupils were discovered whose IQ's ranging between 76 and 89 were no higher than those of pupils in certain of the PS-EMR classes. Their TSI scores exceeded the mean of the special class pupils whose IQ's were in this range.

Implications of the Social Comprehension Deficit:

Among the implications of the widespread and significant deficit in social comprehension by the retarded pupil, to research and rehabilitation, the need for remediation and the importance of an investigation into cause appear most pressing Although some EMR pupils were found to attain social comprehension scores equivalent to those of the NMR, others were found whose TSI performance suggested virtually no comprehension of visual social cues. Although all of the pupils appeared to have benefitted from remedial training, the initially low scoring pupil is seen as particularly in need of remedial assistance.

The discovery of the wide ranges in social comprehension among the retarded, and the overall greater deficit of the institutional pupil, raises the prevention-relevant question as to causes of such gradations in competency. The possibility that factors of personality are in some way related is suggested by the patterns of correlations between pupils' TSI scores and their ratings on the behavioral scales. Specifically, the fact that TSI scores were significantly correlated with ratings on social range, social relationships and peer acceptance and were uncorrelated with ratings on industry and academic skill, suggests the high TSI scorer (the social "comprehender") as being more oriented toward social activities than toward tasks.

A second possibility, suggested by studies of patterns of parent-child interaction, is that factors associated with parental guidance--e.g. failure to convey social information, unrealistically low behavioral expectations, failure to express approval or to otherwise reward childrens' efforts--may be responsible for a portion of the lag in social comprehension.

As a third possibility, it seems reasonable that the quantity and breadth of social experience available to retarded youngsters should be found related to levels of their social comprehension. In this regard, Barker's and Gump's (1964) finding of an inverse relation between size of community or school and frequency of social opportunities for marginal pupils, is worthy of consideration. It was suggested in Chapter 6 that social ecological studies of retarded persons would be of value in providing greater specificity as to their behavioral repertoires, their awareness of the behavioral tolerances of others, and in providing examples of the kinds of situations in which they successfully function and those in which they may not.

Implications of the Remedial Gains in Social Comprehension:

Significant gains in social comprehension were made by pupils using the social perceptual lessons and by those using Fudell's Occupational Educational Units. An implication of these gains is that the comprehension deficit of the retarded is largely remediable through an educative approach. The gains made by pupils using the social perceptual lessons demonstrated that "canned" materials can be effective Although there were differences in the amount of gain between experimental classes (perhaps in part related to teacher liking for the materials or to teacher expertise) gains were attained by nine out of ten of the 127 pupils who used the social perceptual lessons. Of the pupils who used these experimental scripts, 50 percent had gains of 10 or more TSI points when tested at the conclusion of the remedial program; only 12 percent having no gains. Of the 47 pupils who used the contrast lessons, 45 percent had gains of 10 or more TSI points, 31 percent having no gains. In comparison, of the 45 pupils who had no change made in their programs and who were tested and retested over a period of time comparable to the period of special training, 27 percent had gains of 10 or more TSI points and 29 percent had no gains. The teacher presentation of the social perceptual and contrast materials during the experimental trials could not be regarded in any sense as optimal. Teachers were unfamiliar with the lessons and, in order to complete them within the time span, had to present them in rigid sequence without desirable repetition. In normal educational use, the restrictions on time for presentation and the restrictions on teacher-planned variations in use of the lessons would be removed; lessons could be integrated into the year's full program and supplemented, where desirable, by more traditional academic tasks. Under such conditions it seems reasonable to expect greater and more consistent gains in social cue comprehension.

The TSI gain data from the experimental classes were examined for indication that sex, age, or IQ might be important factors in remediability. Sex was found differentially related to gains by pupils who used the first two versions of the social perceptual lessons and by those who used the Occupational Educational Units. In the experimental classes, 63 percent of the boys had gains of 10 or more TSI points, whereas 34 percent of the girls had similar gains. In the contrast classes 32 percent of the boys had gains of 10 or more TSI points, while 56 percent of the girls had similar gains. The consistently greater gains by boys in the experimental classes was thought due to the greater appeal to boys than to girls of the curricula I and II. Trial of curriculum III, rewritten to provide more female relevance, resulted in more equal gains by the two sexes.

The gains by pupils in the experimental classes were examined in relation to three age groupings. Although substantial gains were apparent by pupils in each age grouping, for both sexes the average gain by the youngest (pupils 15 years of age and under) was less than that by the pupils in the next older group (ages 15 to 17) whose TSI gain, in turn, was less than that by the oldest pupils (those 17 and over). Although age apparently had some effect upon rapidity of gain, the investigator's believe that pupil interest in the educational materials and activities was the relevant factor, and that educative treatment could be effective at any age if the materials and lessons were suitably prepared.

The relation of IQ to gains by experimental class pupils was similarly examined. Substantial gains were apparent in relation to each IQ range. For boys, the average TSI gains by those in the two lowest ranges (IQ's of 59 and below below, and IQ's between 60-69) were greater than those by boys in the two higher IQ ranges (70-79, and IQ's over 80). For girls the average TSI gains increased slightly with each of the three lower IQ ranges and slightly decreased for the range of IQ of 80 and above. IQ did not appear as a factor limiting TSI gains.

A question unresolved by the present data concerns the permanence of the experimentally induced TSI gains. Because no followup testing occurred, the stability of the changes cannot be shown. To what extent and under what conditions such gains are maintained is answerable only in terms of more extended study.

Also unresolved is the question of the validity of TSI gains in the sense of their relatedness to other measures of (1) social comprehension, and (2) measures of social behavior. In the current study, an expected relation between increased social comprehension and more adequate behavior was not discerned in the correlations between TSI score gains and gains in teacher ratings. In reviewing the method of TSI scoring it appears possible that gain scores should attain greater validity if the scoring were changed from simple numerical addition to a differential weighting of responses. A series of validation trials would be helpful, in this connection, by providing new criteria for use in item analysis, with the goal of determining whether particular types of responses to certain test items are predictably associated with the retardate's more successful transition to self-maintenance.

Meanwhile, the hypothesis that social comprehension is important to the retarded can be more firmly postulated on the basis of the relationships, other than gain scores, demonstrated in this project's data, and it seems a reasonable conclusion to advocate for the retarded the use of preventive and remedial techniques.

REFERENCES

Baldwin, A.L., Kalhorn, J. & Breese, F.H. Patterns of parent behavior. *Psychol. Monogr.* 1945, 58 (3)

Baldwin, A.L., Kalhorn, J. & Breese, F.H. The appraisal of parent behavior. *Psychol. Monogr.* 1949, 63 (4)

Baldwin, Willie Kate. The social position of the educable mentally retarded child in the regular grades in the public schools. *Except. Child.* 1958, 25, 106-108, 112.

Bandura, A. & Walters, R.H. *Social learning and personality development.* NY: Holt, Rinehart & Winston, Inc. 1967.

Barker, R.G. & Gump, P. *Big school, small school.* Stanford: Stanford University Press, 1964.

Barker, R.G. & Wright, H. *Midwest city and its children: The psychological ecology of an American town.* Evanston, Ill.: Row, Peterson, 1955.

Bass, Daryle I. Assessing manifest coping behavior of mentally retarded institutionalized males with an observational technique. Unpublished Master's thesis, Kansas State College of Pittsburgh, April, 1967.

Beck, S.J. *Rorschach's test.* Vo. 3. NY: Grune & Stratton, 1952.

Beier, E.E., Gorlow, L. & Stacy, C.L. The fantasy life of the mental defective. *Amer. J. ment. Def.* 1951, 55, 582-589.

Brown, R. *Social psychology.* NY: The Free Press, 1965.

Bruner, J.S. On perceptual readiness. In Beardslee, D.C. & Wertheimer, M. /Eds./ *Readings in perception.* Princeton, NJ: D. Van Nostrand, 1958.

Bruner, J.S., Goodnow, Jacqueline J., & Austin, G.A. *A study of thinking.* John Wiley, 1956.

Butler, R.J. Responses of institutionalized mentally retarded children to human and animal pictures. *Amer. J. ment. Def.* 1961, 65 (5) 620 -

Chaffin, J.D. Production rate as a variable in the job success or failure of educable mentally retarded adolescents.1968, *Exceptional Children* (in press)

Chansky, N.M. A study of perceptual training with elementary school underachievers. *J. School Psychol.* 1963, 1, 33-41.

Chansky, N.M. & Taylor, Margaret. Perceptual training with young mental retardates. *Amer. J. ment. Def.* 1964, 68 (4), 460-468.

Clark, D.F. Visual feedback in the social learning of the subnormal. *J. ment. Subnorm.* 1960, 6, 30-39.

Clark, D.F. & Thomson, Stella H.C. An experiment in concept formation in subnormals. *J. ment. Subnorm.* 1963, 9(16) 21-29.

Clark, G.M. Efficacy of an experimental curriculum unit with educable mentally retarded adolescents. Unpublished doctoral dissertation. George Peabody College for Teachers. August 1967.

Clark, G.M. personal communication, 1967 (b).

de Jung, J.E. Effects of rater frames of reference on peer ratings. *J. exp. Educ.* 1964, 33, 121-131.

de Jung, J.E. & Haring, N.G. The measurement of social relations of mentally retarded and young elementary school children using a modification of Syracuse Scales of Social Relations. Final report for Cooperative Research Project Contract #SAE 9177 (Project #1182) 1962.

Denhoff, E. & Robinault, Isabel. *Cerebral palsy and related disorders.* NY: McGraw Hill, 1960.

Deno, Evelyn, Henze, R., Krantz, G. & Borklind, K. Retarded youth: Their school rehabilitation needs. Final report VRA Project RD-681, Minneapolis: Minneapolis Public Schools, 1965.

Doll, E.A. The measurement of social competence: A manual for the Vineland Social Maturity Scale. Minneapolis: Educational Test Bureau, Educational Publishers, 1953.

Driscoll, J.P. The effects of mental retardation on film learning. Los Angeles: University of California. USOE Title VII proj. #365, University Microfilm Pub. No. 63-2853, 1962.

Ebel, R.L. Must all tests be valid? *Amer. Psychologist.* 1961, 16, 640-647.

Edmonson, Barbara, de Jung, J.E. & Leland, H. Social perceptual (nonverbal communication) training of retarded adolescents. *Mental retard J.* 1965,3(5)7-9

Edmonson, Barbara, de Jung, J.E., Leland, H. & Leach, Ethel M. Manual: A Test of Social Inference. Kansas City, Kansas: University of Kansas Medical Center, 1966.

Edmonson, Barbara, Havens, Jane & Carrell, Demaris. Scoring Guide: A Test of Social Inference. Kansas City, Kansas: University of Kansas Medical Center, 1967.

Edmonson, Barbara, Leach, Ethel M., & Leland, H. *Social perceptual training for community living: Pre-vocational units for retarded youth.* Kansas City, Kansas: University of Kansas Medical Center, 1967.

Ekman, P. Communication through non verbal behavior: A source of information about an interpersonal relationship. In Tomkins, S.S. & Izzard, Carroll E. /Eds./ *Affect, cognition, and personality.* NY: Springer Publishing Co., 1965.

Fudell, S.E. Occupational educational units for mentally retarded adolescents. Unpublished doctoral dissertation, University of Texas. December 1963. Order No. 64-37.

Fudell, S.E. & Peck, J.R. *How to hold your job*. NY: John Day Co. 1967.

Gardner, E.F. & Thompson, G.C. *Social relations and morale in small groups*. NY: Appleton-Century-Crofts, 1956.

Garfield, S.L. Abnormal behavior and mental deficiency. In Ellis, N.R. /Ed./ *Handbook of mental deficiency*. NY: McGraw-Hill 1963.

Gellman, W. Vocational rehabilitation: Administration and program development on the local level. In *Vocational training and rehabilitation of exceptional children*. Proc. 1957 Spring Conf. The Wood Schools, Chicago.

Goffman, E. *The presentation of self in everyday life*. Garden City: Doubleday Anchor, 1959.

Goffman, E. *Encounters*. Indianapolis: Bobbs-Merrill, 1961.

Goffman, E. *Behavior in public places*. London: The Free Press of Glencoe, 1963.

Goodman, E.O.,jr. Procedures for the use of sound-recorded dramatic episodes in cultivating the ability to understand social interaction. Unpublished doctoral dissertation. Columbia University, 1962.

Gorlow, L., Butler, A. & Guthrie, G.M. Correlates of self-attitudes of retardates. *Amer. J. ment. Def*. 1963, 67 (Jan). 549-

Guralnick, D. The relationship of personal characteristics and educational experience of mentally retarded persons to their successful utilization of a vocational rehabilitation program. Unpublished doctoral dissertation. New York University, 1963, Order No. 64-256.

Guskin, S. Social psychologies of mental deficiency. In Ellis, N.R. /Ed./ *Handbook of mental deficiency*. NY: McGraw-Hill, 1963.

Hall, E.T. *The silent language*. Greenwich, Conn.: Fawcett Publications, 1959.

Harrison, R.P. Pictic analysis: Toward a vocabulary and syntax for the pictorial code; with research on facial communication. Unpublished doctoral dissertation, Michigan State University, 1964. Order No. 65-6079.

Heber, R. /Ed./ A manual on terminology and classification in mental retardation. *Monog. Supp. Amer. J. ment. Def*. Second edition, 1961.

Harris, D. /Ed./ *Problems in measuring change*. Madison: Univ. Wisc. Press, 1963

Hess, R.D. & Shipman, Virginia C. Early experience and the socialization of cognitive modes in children. Child Develop. 1965, 36 (4) 869-886.

Jackson, Sue L. & Butler, A J. Prediction of successful community placement of institutionalized retardates. Amer. J. ment. Def. 1963, 68 (2) 211-217.

Johnson, G.O. A study of the social position of mentally handicapped children in the regular grades. Amer. J. ment. Def. 1950, 55, 60-89.

Kraft, Elizabeth D. Construction of a movie test to measure the ability to understand others. Unpublished doctoral dissertation. Columbia University, 1962.

Leland, H. Adaptive behavior as related to the treatment of the retarded. J. Music Therapy 1964, 1 (4) 129-131.

Leland, H. & Goldberg, I. Rehabilitation of the institutionalized mentally retarded. Amer. Psychologist 1957, 12 (8) 528-530.

Leland, H. & Smith, D.E. Play therapy with mentally subnormal children. NY: Grune & Stratton, 1965.

Levi, Aurelia. Treatment of a disorder of perception and concept formation in a case of school failure. J. Consult. Psychol. 1965, 29 (4), 289-295.

Lindquist, E.F. Design and analysis of experiments in psychology and education. Boston: Houghton Mifflin Co. 1953.

Lindsley, O.R. Direct measurement and prosthesis of retarded behavior. J. Educ. 1964, 147, 62-81.

Lindsley, O.R. An experiment with parents handling behavior at home. Johnstone Bulletin, 1966, 9, 27-36.

Masland, R.L., Sarason, S.B. & Gladwin, T. Mental subnormality. NY: Basic Books, 1958.

Miller, G.A. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol. Rev. 1956, 63, 81-97.

O'Connor, N. The occupational success of feebleminded adolescents. Occup. Psychol. 1953, 27, 157-164.

Orr, T.B. Social perception in retardates and its modification under stress. Unpublished doctoral dissertation. Indiana University, 1960.

Osgood, C E. Dimensionality of the semantic space for communication via facial expressions. Scand. J. Psychol. 1966, 7, 1-29.

Parnicky, J.J. & Kahn, H. Evaluating and developing vocational potential of institutionalized retarded adolescents. Report of VRA Project 425. Bordentown, NJ: Edward R. Johnstone Training and Research Center, 1963.

Peck, J.R., Stephens, Will Beth & Fooshee, D.K. *The Texas screening battery for subnormals*. Austin: 1804 Raleigh, 1964.

Peterson, L. & Smith, L.L. A comparison of post school adjustment of educable mentally retarded adults with that of adults of normal intelligence. *Excep. Child.* 1960, 26, 404-408.

Rao, C.R. *Advanced statistical methods in biometric research*. NY: Wiley, 1952.

Rosenberg, B.G. & Langer, J. A study of postural-gestural communication. *J. pers. Soc. Psychol.* 1965, 2(4) 593-597.

Ruesch, G. & Kees, W. *Nonverbal communication: Notes on the visual perception of human relations*. Berkeley and Los Angeles: University of California Press, 1956.

Sarason, S.B. The use of the TAT with mentally deficient children. 1. A study of high grade girls. *Amer. J. ment. Def.* 1943, 47, 414-421.

Sarason, S.B. & Gladwin, T. Psychological and cultural problems in mental subnormality: a review of research. *Genet. Psychol. Monogr.* 1958, 57, 3-290.

Scott, W.D. et al. *Personnel Management*. NY: McGraw-Hill, 1941.

Secord, P.F. Facial features and inference processes in interpersonal perception. In Tagiuri, R. & Petrullo, L. /Eds./ *Person perception and interpersonal behavior*. Stanford: Stanford University Press, 1958.

Shellhaas, M. The social and cultural aspects of adaptive behavior. *Parsons State Hospital and Training Center Project News* 1965, 1 (8) 1-4.

Shellhaas, M. Differential social demands in a community survey. Paper presented at convention of AAMD, Boston, Mass., May 1968 (a).

Shellhaas, M. Motion pictures for stimulus presentation: development and uses for opinion-attitude research interviews. *Psychol. Rep.* 1968, 22 689-692 (b).

Siegel, S. *Nonparametric statistics for the behavioral sciences*. NY: McGraw-Hill 1956.

Silverstein, A.B. & Mohan, D.J. Conceptual area analysis of the test performance of mentally retarded adults. *J. abnorm. Soc. Psychol.* 1963, 66 (3) 255-260.

Smith, R.D. Refinement of administration and scoring procedures for the Test of Social Inference. Unpublished doctoral dissertation. University of Oregon, June 1968.

Stables, Jeanne Marie. Self rating of social behavior by junior high special education pupils. Unpublished manuscript. Kansas City, Kansas, Childrens' Rehabilitation Unit, 1966.

Stephens, W.E. A comparison of the performance of normal and subnormal boys on the structural categorization tasks. Except. Child. 1964, 30 (7).

Stephens, W.E. Category usage by normal and mentally retarded boys. Child Develop. 1966, 37 (2) 355-361.

Taylor, J.B. Personality and ability in the lower intellectual range: A study of assessment methods. Final report VRA RD-603. 1960-1963. Seattle, Wash.: Inst. for Sociological Research, University of Washington, 1964.

Ullman, C.A. Teacher, peers and tests as predictors of adjustment. J. educ. Psychol. 1957, 48 (5) 257-267.

Vernon, M.D. The functions of schemata in perceiving. Psychol. Rev. 1955, 62 (3).

Voelker, P.H. The value of certain selected factors in predicting early post-school employment for white educable mentally retarded males. Unpublished doctoral dissertation. University of Michigan, 1962. Order #63-469, DA 1962/1963 XXIII, 3243.

Wallin, J.E.W. Mental deficiency; in relation to problems of genesis, social and occupational consequences, utilization, control, and prevention. J. clin. Psychol. 1956.

Windle, C. Prognosis of mentally subnormals. Monog. Supp. to Amer. J. ment. Def. 1962, 66 (5).

Wunsch, W.L. The first complete tabulation of the Rhode Island mental deficiency register. Amer. J. ment. Def. 1951, 55, 293-312.

Test of Social Inference
Experimental Edition (J 35)

Appendix A

Total Scores

Inf	Error

Code_____ Female_____ Male_____ School____________________ Test _____ Retest_____ Time begin _________________ Time end _________________

Subject ____________________ Birthdate _______________ Date__________ Examiner _________________

Examiner impression of subject:
(physical handicap, attitude toward test, etc.)

Questions	Score	Scorable Responses (Inference)	Enumeration and other Non-scorable Responses	Errors
Practice Picture: Snowman				
A. "What's happening here?" "Who is she?" "What makes you think so?"		Boy/kids put it there	Sees snowman in freezer Mother-lady-maid	
B. "How did it get there?" "What makes you think so?"				
Note: If somebody, ask, "Who?"				
C. "How does she feel about it?"		Surprised-startled-angry-shocked-terrible		Scared Frightened
D. "How does it all go together?" "Can you tell me the story?" This is asked even though subject has given much information in response to the above questions.				

Ⓒ 1966 - Edmonson, Leland, de Jung & Leach, Kansas City, Ks.

Inf	E.

Social Adequacy Criteria and Rating Category Descriptions for Five Teacher Behavior Rating Scales

Social Range: At ease in diverse situations; participant in a broad array of situations.

Very Poor	Poor	Fair	Good	Very Good
Never at ease; restricted; "shrunk": resists any change	Timid or hesitant about new situations	Tolerates a fair range. Is not afraid to try some new activities	Comfortable in many situations. Tries out new activities	At ease anywhere, broad activities

Social Relationship: The ability to get along, fraternize, integrate, converse with peers

Very Poor	Poor	Fair	Good	Very Good
Has problems with most, while not having close relationship with any peer	Mixes well with select few; has open conflict with some; keeps to self	Self sufficient. Not often very friendly; not often unfriendly	Generally friendly; initiates some friendly acts	Mixes well with whole group

Peer Acceptance: Is accepted, integrated, included in conversations and social activities. Is sought out and liked.

Very Poor	Poor	Fair	Good	Very Good
Is rejected by almost everyone	Ignored by most. Maybe tolerated or included in very limited activities	Has some acceptance by a few peers. Included in limited activities	Liked by most peers	Liked and sought by almost everyone. Center of most activities

Social Invisibility: Quality of social behavior enabling subject to fit into situations acceptably.

Very Poor	Poor	Fair	Good	Very Good
Behavior always conspicuously inappropriate, unattractive, or discourteous	Often socially conspicuous by inappropriate or discourteous behavior	Some noticeable mannerisms or behaviors, but generally appropriate. At times conspicuous	Usually well mannered and appropriately behaving	Blends in with groups and activities smoothly. Behavior entirely appropriate and mannered.

Criteria[1] and Rating Category Descriptions for Additional Teacher Behavior Rating Scales (T-BR)

Appearance: Shoes tied, clothing neatly arranged, hair groomed, cleanliness of person and clothing, appropriate clothing, proper make-up (girls), clean shaven (boys).

Very Poor	Poor	Fair	Good	Very Good
Below average on most criteria	Below average on one or more criteria	Average	Above average on one or more criteria	Above average on most criteria

Industry: Ability of student to stay with assigned task: to work with a minimum of conversation, wandering and wasting time. Works spontaneously, with enthusiasm and initiative. Completes tasks.

Very Poor	Poor	Fair	Good	Very Good
Wastes a great deal of time wandering about, usually engages in conversation	Wastes time, frequently wanders about, engages in conversation	Wastes little time. Performs tasks with some interest	Cooperative,shows willingness and industry	Assumes responsibility for completion of tasks. Initiative and interest

Calmness: Emotional control; ability to accept changes in situation without becoming upset; being able to take directions, suggestions, reprimands without losing temper or showing emotional outbursts.

Very Poor	Poor	Fair	Good	Very Good
Generally out of control; in tense mood	Loses control very easily	Generally shows calm attitude to situation	Calm in most situations	Calm in all situations

Academic Skill: Quality of production. Work is done in an acceptable manner.

Very Poor	Poor	Fair	Good	Very Good
Workmanship very poor. Has to be done over often	Will at times do a fair job. Usually sloppy & needs coaching	Work sometimes needs re-doing	Does a good job most of the time	Almost always does good work

---

[1] These scale criteria are from Parnicky and Kahn (1963).

Attentiveness:[1] The ability to respond completely to the task or to the person speaking; ability to keep one's attention focused.

Very Poor	Poor	Fair	Good	Very Good
Attention span so short as to be negligible regardless of task	Difficult to hold attention in most situations	Will pay attention to some situations, or attends with outside help	Usually pays close attention. Some improvement possible	Always pays close attention to directions and work

---

[1] Adapted from Parnicky and Kahn (1963)

Outside Rater Behavior Rating Scales (O-BR)

Rater Instructions

1. Read the criteria to be used in evaluating pupil performance.

2. Call upon your experience with retarded youth of similar ages (___to___) and try to recollect one individual as the most outstanding in respect to the criterion description. Write his or her name in the box marked "most." Next, recall the poorest individual you can remember with respect to the criterion and enter his or her name in the "least" box. Next, fill in the "middle" box with the name of a retarded youth who fits the description "middle." Then complete the frame of reference by writing in names of two boys or girls you recall as midway between the poorest and the middle and between middle and most outstanding.

3. Evaluate each of the listed pupils with respect to criterion. Make a check mark at the segment of the line adjacent each ratee that indicates where he stands in relation to the frame of reference.

4. When you finish, check back to make sure that each pupil has been given a rating.

1. Appearance: Dress is appropriate to situations; is becoming, fits well, is clean, pressed, and mended. Hands, face and hair clean and well groomed. No body odor. Proper make-up (girls). Shoes clean, in good repair and appropriate to costume and situation; tied.

Frame of Reference	least Slovenly.Below average on most.	between Below average on one or more criteria	middle Average	between Above average on one or more criteria	most Above average on most criteria
Pupils					
1.__________	__________	__________	__________	__________	__________
2.__________	__________	__________	__________	__________	__________
Etc.__________	__________	__________	__________	__________	__________

2. Industry: Begins work promptly and willingly; works with minimum of wasted time, conversation, day dreaming, etc.; persists at assigned tasks to completion.

least	between	middle	between	most
Wastes a great deal of time	Wastes time frequently. May wander, converse,etc.	Wastes little time	Cooperative;shows willingness and industry	Assumes responsibility for tasks. Initiative and interest

3. Social Acceptance: Is definitely sought out and much liked by peers. Others continually drawn toward him or her. Center of activities.

least	between	middle	between	most
No clase relationships. Rejected	Mixes with a few, conflict with some; mostly keeps to self	Self sufficient. Not often unfriendly; not often very friendly	Generally friendly; Initiates some friendly acts.	Mixes well

4. Social Range: At ease anywhere in diverse situations, settings and groups; participant in a broad array of situations.

least	between	middle	between	most
Never at ease, restricted; "shrunk"	Timid or hesitant about many situations	Tolerates a fair range	Comfortable in many situations. Tries some new activities	At ease most anywhere; broad activities

5. Calmness: Emotional control, ability to accept changes in situations without becoming upset; able to take suggestions and reprimands without losing temper or showing emotional outbursts. Able to take kidding.

least	between	middle	between	most
Extremely touchy, irritable; oversensitive; frequent outbursts	Touchy,irritable, defensive with peers & adults	Reasonably calm but may respond better to peers than to adults or vice versa	Usually responds constructively to adults & peers	Always calm; responds well to correction & peer generated problems

Appendix C (cont'd)

6. Attentiveness and receptivity: Responds completely to task or person speaking; listens and retains and follows instructions. Pays close attention to work.

least	between	middle	between	most
Negligible attention span regardless of task	Often difficult to hold attention and foster retention	Will pay attention when interested, or with outside help. Some retention	Tries to be attentive & retentive. Some improvement possible	Pays close attention to directions and work

7. Quality of Production: Measures up to high standards of work output.

least	between	middle	between	most
Workmanship very poor	Usually a little sloppy. Needs coaching.	Work a little uneven. Sometimes needs re-doing	Usually does a good job	Almost always does good work

8. Social Invisibility: Blends in with groups and activities smoothly. Behavior entirely appropriate and mannered.

least	between	middle	between	most
Behavior always conspicuously inappropriate, unattractive or discourteous	Often socially conspicuously inappropriate or discourteous behavior	Some noticeable mannerisms but generally appropriate	Usually fits in	Blends with groups and situations smoothly. Appropriate and mannered

Remember to encourage the less verbal or slower pupils to respond at times when questions are at their level--so all pupils are encouraged to attend, think, and respond

PLACES SOMETIMES TELL US WHAT TO DO*

Objectives: To call attention to settings which communicate the terms on which admission is possible

Materials
Slide projector
Screen
Set Slides (III-A)

Vocabulary:
admission requirements
badge
bullseye
depot
emergency room
identification
injured
knuckles
museum
night chain
no trespassing
on display
padlock
peep hole
privacy
protection
public library
public park
public place
station
store hours
terminal
undesirable
visitors hours
welcome

We have been learning about the many kinds of signals that tell us what we are supposed to do. /Teacher may wish to review accident preventing signals and way-finding signals./

This week we will look at pictures of public places. As we look at the
pictures, we will often find that we can know what we are supposed to do. As
(1) an example /Slide: Fence/. We can read what a fence or a wall tells us to
do. It tells us to "Stay out."

(2) /Slide: Door with padlock/ There is a padlock on the door. What about this one? What does it tell us to do? ( ) It's another way to say "Keep out."

(3) /Slide: High fence and soldier guard/ What is this? (A fence and guard.) What does it tell us? ("Stay out." Do you think it politely says, "Please stay out?" ( ) No, it says "Keep out. We mean business!" How can we know? What makes us think so? (The guard, gun, barbed wire, and very high fence.)

There are different ways of telling people to keep out--stop--telling them not to enter. When we look at something and ask ourself what it is trying to tell us, sometimes we can read the message

---

* © 1967, Edmonson, Leach & Leland, Kansas City, Kansas

(4) As an example, this fence looks different from the last one we saw. /Slide: Picket fence/ It doesn't say, "Everybody keep out." It is a very low fence and is probably meant to keep children and dogs from running through the yard and garden. I think this says, "Come in where the gate is; stay on the sidewalk." What do you think? ( )

Why do people build walls and fences? (/Encourage discussion. What would it be like if there were no walls around our classroom? What would it be like if there were no walls around where we sleep? toilet? bathe? etc.?/) What are some reasons why we build walls around where we live? ( ) What are some reasons why we build walls around where people work? ( ) /Teacher may relate responses to desire for privacy, protection; may discuss meaning of both./

We can understand why there are walls or fences around most places. It is because we need some control over who comes in or who goes out. We make openings in the walls as a place where people can come in, but we usually put a door in the opening. Why? (It can be closed to keep out noise, confusion, undesirable people, or regulate the traffic of people coming in and out.) Why should we be able to lock doors? (Because we need privacy and protection.) People cannot walk in and out of every place just because they would like to. Before we go in, we should find out if others want us to come in--are we welcome?

(5) /Slide: Home entrance/ That is why there is usually a doorbell or a door knocker on a house. Sometimes there isn't a doorbell or a knocker, then what do we do? (Knock) Yes, a good thing to do is to knock on the door, or the wall, with our knuckles. /Discussion may continue and be made more specific, with reference to when one knocks and when one might call out "Is anyone in?" - etc.. In many situations it is not courteous to knock and walk in without an invitation./

People do not have to let other people come into their homes. Sometimes they have a window in the door so they can look out and see who is there. /May discuss what pupils' doors are like at home--a night chain, a bullseye, etc.. Then discuss what parents do when the doorbell rings, or when there is a knock at the door. May discuss what cues people might look for--someone they know, a child, a uniformed person, etc.--or what they might ask./

A home is a private place. That means we have the right to protect ourselves and what we own. We can let some people in and keep others out. There are some people who come to our door we let come in. Let's think about how we decide on the people we let in and the people we don't let in. /Encourage discussion. Good responses might include: We let in friends, at reasonable times./ Most people are probably good people and would not disturb us or hurt us, but some people might try to take our things from us or hurt us, so we should be careful about opening the door and letting people come in. We let young people or children visit our homes to play. We let neighbors come to our homes to visit. We need to be careful about young people or grownups whom we do not know very well.

Can we think of some other places besides our homes that are private?
(6) ([Teacher accepts reasonable responses.]) [Slide: No trespassing] A yard
or a lot or a farm is private property. The owner usually does not want
everybody to walk across the grounds. He might put up a sign like this to
(7) remind people it is private property. [Slide: Keep Off the Grass] He
(8) might put up a sign like this. [Slide: Keep Out] [Discuss other restrictions re privacy and property.]*

The grounds around a factory are private property. It is the property
of the people who own the factory. They may put up a fence to keep people
(9) out who do not work there. [Slide: Security factory] This factory has a
fence and a guard house or a security check. The guard looks at the workers
(10) as they come through the gate [Slide: Close up] and they must be wearing
(11) a special badge with their picture on it or he will not let them in. [Slide:
Badge] They must identify themselves.

What about a hospital. Would you call it a private or a public place?
[Teacher encourages discussion.] It is partly both. There is one part--the
emergency room--where anyone can go if he is sick or injured. It is always
open. But most of the hospital building is private. The people who are in
bed in the hospital rooms and wards are supposed to rest and not be disturbed.
(12) Some of the people are well enough to have visitors. [Slide: Visitor's
hours] What does this tell us? ( ) There are times when friends and
family can visit someone in the hospital. Young children are not supposed
to visit people in hospitals. To be allowed to visit in this hospital you
must be 14 years old.**

We know something about what we are supposed to do when we are in a hospital, so as not to disturb people. What is it? (Good responses are accepted.])

A store is a private place. The people who own the store pay for everything that is in the store. All of the merchandise in the store is private property. They will allow us to buy their merchandise. They will allow us, at reasonable times, to come in and look at the merchandise they have on display. If we do not see what we want, we can ask if they have it. A_____, is it good behavior for customers to go behind the counters? ( ) To look in drawers? ( ) B_____, to take things that belong to others (stores)? ( ) To do anything that disturbs other people in the store? ( )

Some places are public places. A church is a public place. A train station or a bus depot is a public place. Anyone can come in. A museum or a public library or a public park is a public place, and anyone can come in if they follow certain rules. [Discuss expected behavior.]

A few public places are always "open." By that I mean the door is never locked. The bus depot and the air terminal and the train station would be

---

* Possibilities: Things we buy or receive as gifts are our private property. We have the right to keep them or to lend them or not lend them to others. Other people are not behaving as they should when they look at things we haven't shown them or when they borrow things without permission.

** Regulation differs from hospital to hospital

open all the time in a big town because buses and trains and airplanes are
coming and going all day and night. A public park might stay open. In some
towns there might be a fence around the park with a gate. They might close
a gate and lock it at night. Most public places are closed at certain times,
(13) especially at night. /Slide: Hours, days/ This is a sign on a door at a
supermarket. What does it tell us? ( ) It tells us when the store is
(14) open. /Slide: Hours, days/ This is a sign on a door of a bank. What
does it tell us? ( ) It tells us when the bank is open. /May discuss
places known to pupils that remain open or that are closed at certain times./

/Discussion of private and public can continue if fruitful./

Optional and Supplemental Activities

Discuss places visited during the week. Decide whether private or public, and what pupils know about the "admission requirements."

Discuss possible reasons for the hours and days of admission and no admission.

Discuss personal experiences with borrowing or lending of personal possessions, in particular the kinds of problems that occurred and how they might be prevented or resolved.

Appendix E

LESSON TOPICS AND INTRODUCTION TO COMPETING CURRICULUM USED IN TRIALS 2 AND 3

Twelve Units in Occupational Education For Secondary School Mentally Retarded Youth

Introduction

One of the primary purposes of the secondary school educable mentally retarded program is to prepare students for gainful employment. The objectives of the twelve units is to instill in these students values and attitudes that are basic and necessary for the success of any worker. There will be no attempt to train these students for any specific trades or skills but rather an attempt to give them attitudes and values that are personality oriented. The following list of the twelve units will clarify the objectives.

1. Why Should We Learn About Jobs?
2. Getting Along With Our Fellow Workers
3. Are You Willing to Work and Do Your Best?
4. Following Directions and Finishing Your Work
5. Being On Time
6. Are You Reliable? Can We Depend On You?
7. Honesty Is the Best
8. You Must Work Harder and Produce More if You Want A Raise in Salary or Grades
9. 50 Ways to Lose Your Job
10. Cooperation is the Basis of Success
11. Your Attitude Means Success or Failure
12. Going for That First Job Interview

All units are being written on a second grade reading level. Any student who has successfully completed a second grade basal reader is presumed to be reading at second grade level. The reading vocabulary is made up of all the words in the Dolch Basic Word List, and the Ginn and Scott Foresman basal readers up to and including the second grade. When any new vocabulary is introduced it will be taught as new material. This reading level should be within the ability of most educable mentally retarded secondary school students. It is assumed that students can add prefixes and suffixes and combine words known to them.

It is the aim of the investigator to determine if values and attitudes that are necessary for success in the socio-economic world can be taught in a classroom. The contents of the units were determined on the basis of forty-five interviews with employers or prospective employers of unskilled labor. The contents are not imaginative. The employers interviewed had retarded workers in their employ or could hire retarded workers. These units are a culmination of the thinking of forty-five employers who actually work with retarded people, or "normal" people everyday. These units are as practical as possible. Some Teaching Suggestions:

Appendix E (cont'd)

1. Always take the non-readers in one group and read the material to them. One of the better reading students can do this task for you. Do not exclude them because of lack of ability to read.

2. Be enthusiastic. These contents and their continual repetition will at times bore you. Do not let this affect your enthusiasm. As you go, so goes the class. The technique of a constant repetition of facts is needed for retarded children.

3. Always explain directions carefully, especially on Day 5, test day.

4. Try to carry on the unit's contents in daily class life. If at all possible, make it a 24-hour living experience. Carry over these principles into other subject matter. Try to ingrain these good habits while the students are in school.

5. The teacher will notice that much of this material is presented in a question-answer approach. This technique has not been an often used one with retarded students. The writer has found it to be an excellent technique for presenting new material. One note of caution. Unless the teacher uses a large amount of showmanship and enthusiasm this question-answer technique will fail. Frankly, a third-degree type of questioning is best. You will have to probe for an answer as you want it. Often times you will have to skip answers until you receive the answer you want. When you receive the correct answer as shown in the unit, put it on the board. The students will think that these answers are theirs. You can usually say, "But you said this, not me," in review or further work.

Appendix F

Movies and Filmstrips Used with Competing Curriculum in Trial 2

Unit I	None	
Unit II	"Words of Courtesy" (McG-H)	
Unit III	"How to Keep a Job" (Cor)	
Unit IV	"Ways to Good Habits" (Cor)	7641 "School Safety" (EBF)
Unit V	None	
Unit VI	"Am I Trustworthy" (Cor)	8625 "Promises Are Made To Keep" (EBF)
Unit VII	None	8495 "One Kind of Bravery" (EBF)
Unit VIII	"Your Earning Power" (Cor)	
Unit XI	None	
Unit X	"Office Teamwork" (EBF)	
Unit XI	"Act Your Age" (Cor) "Griper" (McG-H)	
Unit XII	"Personal Qualities for Job Success" (Cor)	1171 "The Job Interview" (EGH)

Cor - Coronet
EBF - Encyclopedia Britannica Films
EGH - Eye Gate House
McG-H - McGraw-Hill

Appendix F (cont'd)

Notes for Teachers in re Movies and Filmstrips Used in Trial 2

Unit I - None

Unit II - "Words of Courtesy" - 11 min. (McG-H)

This film deals with children 12 or 13 in the home, school and community. It contrasts good and wrong ways of asking for things. It emphasizes the importance of "thank you", thoughtfulness, and a friendly attitude to others. Teacher would have to help relate material to vocational setting.

Unit III - "How to Keep a Job" - 11 min. (Cor)

Teacher will have to read ad and employment termination notice to class. Some of the vocabulary is difficult and will have to be interpreted. Film illustrates interview, timeclock, and work in shipping room. Contrasts work habits of two fellows as to initiative, being on time, loyalty, dependability and cooperation.

Unit IV - "Ways to Good Habits" - 11 min. (Cor)

Starts with contrast between good and bad baseball batter. Then contrasts good and bad habits. Pupil is encouraged to practice building good new habits. Should be related to vocational situation.

Filmstrip - 7641 - "School Safety" (EBF)

Teacher may have to read written material to students. Teacher will have to relate safety rules from school to vocational situation.

Unit V - None

Unit VI - "Am I Trustworthy" - 11 min. (Cor)

This film deals with home and school situations in which a boy is helped to understand the meaning of dependability which is made up of being on time, returning things you borrow, doing jobs you undertake. Vocabulary is not too difficult, but again material has to be specifically related to Unit VI.

Filmstrip - "Promises Are Made To Keep" (EBF)

This tells of one girl scout who didn't keep her promise and how she realizes that dependability is important. Written material will have to be read. This material will have to be directly related to Unit VI.

Unit VII - "One Kind of Bravery" - Filmstrip. (EBF)

Written material will have to be read to students. Young boy breaks a window and finally confesses. This is a rather juvenile film. Teacher will have to relate this to the importance of honesty in a vocational situation.

Unit VIII - "Your Earning Power" - 11 min. (Cor)

Vocabulary and some of the concepts illustrated are difficult; i.e., vocational opportunity, economic society, producing and exchanging goods. It is doubtful that these concepts should be fully explained. However, the film does show the relationship between kind of work done, personal qualifications, production, and amount you earn. Jobs illustrated are stock boy, sales clerk, personnel man, delivery man.

Unit IX - None

Unit X - "Office Teamwork" - 11 min. (EBF)

Film is too complicated and setting is unrealistic for retarded. However, there are good concepts illustrated - "Everyone working together puts us on top." Also a griper is shown who constantly complains and this is corrected. Teacher will have to skim over difficult concepts and emphasize understandable concepts related to "vocational teamwork."

Unit XI - "Act Your Age" (Cor)

This film deals with emotional maturity in a school setting. It would have to be related to a vocational setting. A boy can't do a problem and scratches initials in a desk. He has to see principal. The film deals with infantile behavior reactions. Best part of the film is establishment of rating scale on behavior:

"How old am I when (1) I don't get what I want?
" " " " (2) Someone pokes fun at me?
" " " " (3) I want attention, etc.?"

"Griper" - 11 min. (McG-H)

Again the setting is home and school and would have to be related to vocational setting. Film illustrates conscience as a "ghost" stating problem of the griper which might be difficult for retarded to understand. Film wants individual to see himself as others see him. Ends with "Do you know a griper? What do you think?"

Unit XII - "Personal Qualities for Job Success" - 11 min. (Cor)

Vocabulary is difficult. Illustrates way to success and making a good impression in the job interview. Shows machine shop and an office. Tells of necessity for good grooming, being courteous, having an interest in people, and having good work habits.

"The Job Interview" - Filmstrip (EGH)

This is good. Teacher should read written material. Each picture might serve as a focal point for discussion in terms of do's and don'ts in applying for a job.